NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

INVESTIGATING A NEW APPROACH TO SPACE-BASED INFORMATION NETWORKS

by

Joseph C. Horvath
Richard L. Tackett
Cameron R. Yaste

September 2012

Thesis Advisor:	Alex Bordetsky
Second Reader:	Daniel Bursch

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704–0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Investigating a New Approach to Space-Based Information Networks			**5. FUNDING NUMBERS**
6. AUTHOR(S) Joseph C. Horvath, Richard L. Tackett, Cameron R. Yaste			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ___N/A____.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Just as terrestrial networks have shifted to routed, message centric networks, the satellite networking regime is beginning to transition in the same manner. Based on established networking theories and unique orbital models, a satellite based communications network architecture is modeled that would fulfill the requirements of capacity, coverage and logistics for current and future military needs. Specifically, various satellite constellations and network characteristics are examined with a focus on how to enhance current methods of routing data in this environment. The authors explore how Social Networking theories can be applied to the routing decision making process in the space environment. Utilizing aspects of adaptation and the weak-tie relationship, an adaptive hybrid routing protocol model is proposed leveraging the predictability of orbital mechanics to increase performance and decrease network overhead.

14. SUBJECT TERMS Satellite, Network, Routing Protocols			**15. NUMBER OF PAGES** 183
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01–280–5500 Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

INVESTIGATING A NEW APPROACH FOR SPACE-BASED INFORMATION NETWORKS

Joseph C. Horvath
Major, United States Marine Corps
B.S., University of Minnesota 2001

Richard L. Tackett
Captain, United States Marine Corps
B.S., The Citadel 2005

Cameron R. Yaste
Lieutenant, United States Navy
B.A., The Citadel 2006

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SPACE SYSTEMS OPERATIONS

from the

NAVAL POSTGRADUATE SCHOOL
September 2012

Author: Joseph C. Horvath
Richard L. Tackett
Cameron R. Yaste

Approved by: Dr. Alex Bordetsky
Thesis Advisor

Daniel Bursch
Second Reader

Rudy Panholzer
Chair, Department of Space Systems Academic Group

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Just as terrestrial networks have shifted to routed, message centric networks, the satellite networking regime is beginning to transition in the same manner. Based on established networking theories and unique orbital models, a satellite based communications network architecture is modeled that would fulfill the requirements of capacity, coverage and logistics for current and future military needs. Specifically, various satellite constellations and network characteristics are examined with a focus on how to enhance current methods of routing data in this environment. The authors explore how Social Networking theories can be applied to the routing decision making process in the space environment. Utilizing aspects of adaptation and the weak-tie relationship, an adaptive hybrid routing protocol model is proposed leveraging the predictability of orbital mechanics to increase performance and decrease network overhead.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ADOV	Ad-hoc On-Demand Distance Vector
AEHF	Advanced Extremely High Frequency
AGI	Analytical Graphics Incorporated
BER	Bit Error Rate
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CENETIX	Center for Network Innovation and Experimentation
CLEO	Cisco Router in Low Earth Orbit
COE	Classical Orbital Elements
COM	Common Object Model
CSMA	Carrier Sense Multiple Access
DSCS	Defense Satellite Communication System
DTN	Delay Tolerant Networks
DYMO	Dynamic MANET ON-Demand
EHF	Extremely High Frequency
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FLTSATCOM	Fleet Satellite Communications System
FOV	Field of View
GBS	Global Broadcast System
GEO	Geosynchronous Orbit
GUI	Graphical User Interface
HEO	Highly Elliptical Orbit
HRPS	Hybrid Routing Protocol for Space
IETF	Internet Engineering Task Force
IGO	Inter-Governmental Organization
IP	Internet Protocol
IRIS	Internet Router in Space
ISL	Inter-Satellite Link
LEO	Low Earth Orbit
LLC	Logical Link Control
LSACK	Link State Acknowledgement
LSREQ	Link State Request

MAC	Media Access Control
MANET	Mobile Ad-hoc Network
MEO	Medium Earth Orbit
MIB	Management Information Base
MILSATCOM	Military Satellite Communications
MILSTAR	Military Strategic and Tactical Relay
MPR	Multi Point Relays
MUOS	Mobile User Objective System
OLSR	Optimized Link state Routing
OSI	Open Systems Interconnect
OSPFv2	Open Shortest Path First Version 2
PSD	Predetermined Satellite Database
PSMA	Phase Shift Multiple Access
RERR	Route Error
RF	Radio Frequency
RFC	Request For Comment
RREQ	Route Request
SHF	Super High Frequency
SNMP	Simple Network Management Protocol
SNOC	Satellite Network Operations Center
SNR	Signal to Noise Ratio
SSAR	Socially Selfish Aware Routing
STAR	Source Tree Adaptive Routing
STK	Satellite Tool Kit
SWAP	Size Weight and Power
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TL	Trust Level
TT&C	Telemetry Tracking and Communications
UAV	Unmanned Aerial Vehicle
UFO	UHF Follow On
UHF	Ultra High Frequency
VBR	Variable Bit Rate
WGS	Wideband Global System

ACKNOWLEDGMENTS

The Authors would like to thank first and foremost Cisco Systems, and specifically Erik Watco and his team for their support of this project. Without their help and guidance, we would not have been able to accomplish so much. Also we would like to thank Marty Paulsen of Scalable Networks Inc for his generous support of our efforts to obtain software licenses. In addition we would like to thank AGI and Scalable Networks for the fine modeling software products that made this project possible.

We thank Dr. Alex Bordetsky for his unwavering support of our vision, no matter how fuzzy that vision got, and for never failing to trust that we were in fact on to something, for his studious editing, and for his generally spirited support, without which, we would have given up hope long before. We would like to thank the Space Systems Academic Group for challenging us to always look for new ways of thinking and stretching our creative imaginations in order to create new ideas as well as for studiously teaching us, not always an easy task, to understand the vagaries of that esoteric realm aptly named Space.

Finally, we would like to thank our families and friends, who tolerated our late night study sessions, our frustrating moments of panic, and our often exhausted demeanors, without you, it would not have been worth doing, let along possible.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND AND OBJECTIVES

The warfighter of today requires exceedingly more information than that of generations past. Smart weapon technologies, unmanned sensors and ever-changing rules of engagement incur a new challenge in moving information around the battlefield. Current technologies are capable of relaying relatively large flows of data to fixed sites and ships around the world. Unfortunately, there exists a gap in closing the last mile to the displaced warfighter. The problem that faces future military communication networks is three fold: Capacity, Coverage and Logistics.

1. Capacity

Current military communication systems are limited in capacity due to the assets available and the method used to transmit and receive data. This is especially the case for tactical units that do not have the same access as large fixed units. In essence, there are three aspects of capacity that are a requirement for future military operations: Bandwidth, Data Rate and Timeliness. The goal of any future system should be to maximize bandwidth and data rate while minimizing any delays to the end user.

2. Coverage

Due to the global nature of the current threats, a requirement exists for worldwide network coverage. Currently this is accomplished through major hubs around the world, but does not extend into more austere environments.

3. Logistics

There are currently a plethora of network and communications systems, many of them unique to various services, communities and vehicles. This method of ad-hoc acquisitions fails to realize the benefits of a shared network

model. Future systems need to be based on a generic architecture that can be designed to suit the needs of the user.

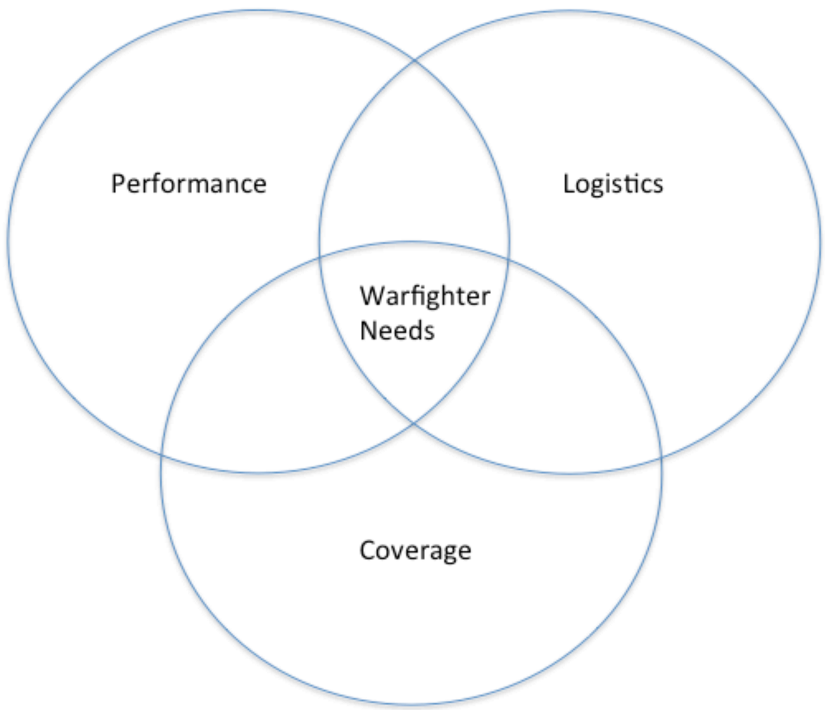

Figure 1. Future Network Tradespace

Current military networks are highly dependent on satellite communications (satcom) to provide backhaul from austere locations. Whether due to insufficient infrastructure or enemy threats, SATCOM is a requirement for any long term military operation. As well, with the current focus on distributed operations, unmanned platforms and unattended sensors, a generic space based network could provide the performance and flexibility that the future requires. In order to accomplish this, a new paradigm for space networks including full orbital routing is required.

B. RESEARCH QUESTION

The goal of this thesis is to develop, model and analyze a space based communications network capable of providing the performance, coverage and logistics required of current and future military needs. While research has been conducted into satellite networking and the future of military communications, a new model is proposed leveraging space based routing technologies, in order to solve this problem. Based on our analysis of established networking theories and unique orbital models, a satellite based communications network architecture is modeled that would fulfill these requirements. Specifically, various satellite constellations and network characteristics are examined with a focus on how to enhance current methods of routing data in this environment.

C. SCOPE AND LIMITATIONS

The nature of this research was to examine space based networks with respect to the network topology (satellite constellations) and packet routing point of view. In order to model and test these areas, the physical layer of the network was restricted. This was accomplished by standardizing the physical layer such that connection between nodes via RF links was always possible when nodes were in view. Due to this optimization, a different approach would be required to examine the physical links in future work.

Currently only one satellite on orbit has onboard routing capability as discussed later. While attempts were made to access this capability, real world testing was not possible. Due to this, the research was limited to computer based modeling and simulation. This modeling was done based on current systems and methods, but evaluation was limited to the available resources in the modeling software. A second limiting factor was the required computing time based on available assets. Due to this factor, the simulation time was chosen such that the desired data could be obtained within a reasonable time frame.

D. METHODOLOGY

In order to accomplish the goals, the project was broken into four phases of research. The first phase was to examine the current methods used in ground and space networks and how a routed space based network could accomplish them. The study also examined previous attempts at developing space based networks with a focus on what benefits were found to exist over terrestrial networks and why these projects failed to come to fruition. The second phase was to examine the types of satellite constellations that could be used to accomplish these requirements. This led to four orbital models used to explore the trade-offs of the physical limitations of space based communication nodes. Using the four constellation models and the communication link metrics developed, phase three used software emulators to test network topologies and protocols in the lab environment. The final phase was to develop a new network routing protocol capable of more efficiently handling the requirement of a space based model.

E. ORGANIZATION OF STUDY

The initial chapters seek to give the reader an essential background on previous and current satcom development. This includes a review of the physical aspects of satellite orbits and constellation development. Following these, the satellite network test bed design and methodology will be explored from both a satellite constellation and an ad-hoc space network perspective. The results of these simulations will be detailed with an examination of their application to the problem being researched. Finally, the team will present a new model for space routing and provide final conclusions and recommendations.

II. BACKGROUND

A. TERRESTRIAL TELECOMMUNICATIONS NETWORKS

The evolution of terrestrial telecommunications from basic telegraph and telephones through the Internet and cellular data networks has gone through multiple stages that are being repeated in satellite communications today. The multiple stages are; point to point communications, circuit switched networks, channelization, and finally message centric networks (Clarke 1945). The drive to converge a communications network that is equipment, location, transport, and media agnostic requires certain architectures, regardless of whether that network is to include satellite nodes. The three main requirements that drive the evolution of telecommunications, both terrestrial and satellite, are constantly increasing demand for bandwidth, constant demand for ubiquitous coverage, and the ability to interconnect various devices and platforms into the same network. This coincident demand of performance, logistics, and coverage are what push telecommunications forward. Given a thorough understanding of the evolution of terrestrial telecommunications and the current state of satellite communications, one can predict where and how the next evolution in communications technology will take place (Elbert 2012).

1. Point-to-Point Communications

The very first communications systems designed to operate at great distances were point to point architectures. There were optical signaling systems dating back to the Middle Ages. They were very basic, often able to transmit a single signal in an emergency, over the range of the human eye. The first electrical signaling system was developed in 1809 by a German named Samuel Tomas Von Sommering. His system consisted of 35 glass and acid insulated wires up to a few kilometers in length. Each wire represented one symbol and the signal was sent by applying an electric charge to each wire. At the recipient's end the charge would electrolyze the acid in its corresponding tube, releasing a

stream of hydrogen bubbles next to its associated symbol. The telegraph operator would visually observe the bubbles and could then record the transmitted message. In 1837, Samuel Morse patented the Morse telegraph and within two decades, the continent of North America was connected via telegraph cables (Michaelis 1965).

The point to point architecture first allowed for messages to be transmitted at light speed across vast distances, but the system was incapable of allowing multiple users, incorporating multiple transmission media, or supporting mobile users and by the end of the same century it was obsolete (Michaelis 1965).

2. Circuit Switched Networks

Since the invention of the telephone by Alexander Graham Bell in the mid 1870s, the world has been increasingly connected, and increasingly subject to the network effect. As telephones proliferated and telephone networks crisscrossed the world, competition between competing carriers centered on this network effect as companies strove to attract and retain the highest number of customers. Since customers value the telephone service based on the number of people they are connected to, the company with the highest level of connectedness is given the highest value. This is the network effect, and it drives telephony service to be interconnected (Neuchterlein and Weiser 2005).

By the 1960's, telephone service was ubiquitous. It did not matter what brand of phone a person owned, where they lived, or what telephone carrier they subscribed to, if they wanted to make a phone call they simply dialed the number and a connection was established. This is called network convergence, or the movement toward uniformity. Analog telephony networks converged thanks to the interconnect ability between different carriers and different phones. The telephone system was a circuit switched network, a network in which each exchange of data is preceded by the establishment of a dedicated circuit between two end points. Once the phone call is finished, the circuit is freed up to be used by another user. Circuit switched networks can accommodate cross

media systems and as early as 1946, radio telephones were in operation in the United States. The use of wireless RF links to backhaul into a wired circuit switched telephony system allowed "mobile" phones to enter the market (Elbert 2008).

The problem with circuit switched networks is that they do not efficiently utilize the network capacity. As demand increases for telephone channels, the circuit switched model gets congested and quality of service declines. A circuit switched network cannot respond to changing traffic demands as circuits must be pre-provisioned and a connection established prior to data transfer. The solution to this problem is to encode multiple analog signals digitally onto the same channel. This allows multiple phone calls to use the same phone line (Subramanian 2000).

3. Channelization

The demand for communications capacity increased dramatically and that demand drove the invention and implementation of several methods of sending multiple messages across the same medium in order to maximize radio frequency (RF) utilization. The idea is to separate the channel into smaller pieces and assign a different customer circuit to each of the smaller pieces. The major schemes for accomplishing this include frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA) (Ippolito Jr. 2008).

a. FDMA

Frequency Division Multiple Access is a schema that breaks down a channel into discreet, lower bandwidth channels on a specific frequency of the main channel. The benefits of this system is that each circuit within the system gets some of the RF power all of the time and it requires no synchronization in order to function (Ippolito Jr. 2008).

b. TDMA

Time Division Multiple Access allows each circuit in the system to use the entire bandwidth of the channel but only for a window of time. The time allocation can be static as a polling sequence or dynamic as a demand assigned multiple access. TDMA requires timing synchronization with all stations in the network (Ippolito Jr. 2008).

c. CDMA

Code Division Multiple Access uses spread spectrum techniques to push the message signal down below the carrier noise level and spreads the message out to cover a much wider bandwidth, allowing multiple message signals, each with their own specific code identifying them, to be transmitted across the same carrier (Ippolito Jr. 2008).

Multiple access schemes have allowed communications networks to increase overall capacity and reduce “dead time” on each physical link. They have also been the major enabling technology for the proliferation of cellular telephones by allowing large numbers of users to share bandwidth. As the network effects increase demands for ubiquitous coverage for any and all forms of data, channelization is an inevitable step toward network convergence. Up to this point in the evolution of communications networks, the focus has been on the end to end connection. The defining characteristic of the network has been the circuits formed between end users, or how to better use physical media to connect more people. The network of this type has no cognizance of the different types of messages being sent across them.

4. Message-Centric Networks

The first packet switched network, ARPANET, was developed to support research at four major universities in the United States. The real revolution of ARPANET was the paradigm shift from circuit centric networks to message centric networks. A combined effort between researchers in the U.S. and UK

developed the logical and mathematical framework that enabled the network to evolve from a collection of independent connections into a collection of independent messages. Each message is separated into packets and the transport media used is decided on a per packet basis at each node in the network (Dean 2010).

Packet-switching provides better bandwidth utilization and response times than the traditional circuit-switching technology used for telephony, particularly on resource-limited interconnection links. Packet switching also allows for different pieces of a single message to be routed along different physical paths, making the entire network very robust. Packet switching is also transport agnostic. A message from a node connected wirelessly via a 2.4GHz Wi-Fi link can easily communicate with a node connected via optical fiber thousands of miles away, without having to engineer an interface mechanism to support the change of transport medium. Packet switching also requires very little pre-provisioning of services since end to end quality of service is handled on an aggregated basis (Dean 2010).

B. SATELLITE COMMUNICATIONS

While many people had recognized the existence of a geosynchronous orbit, the original idea for using such an orbit for communications relay satellites is attributed to the science fiction author, Arthur C. Clark (Clarke 1945). His article in *Wireless World* magazine in 1945 laid out the detailed plan for a constellation of three satellites to provide a worldwide network for communications. The original concept was to provide satellite based relays to connect continents. Arthur C. Clarke understood that the technology he envisioned was at least a few years from maturation, and indeed the first operational geosynchronous spacecraft was not launched until the early 1960’s (Clarke 1945).

1. Point-to-Point Communication

During the early days of the space race between the United States and the USSR, the GEO satellite idea was thought to be too speculative and instead efforts were focused on the low earth orbit (LEO) passive reflector and active repeater programs. NASA pursued several permutations of the LEO communications satellite including programs like Echo, Advent, Telstar, Courier, and Relay in the hopes of competing with terrestrial wire. The feeling was that the propagation delay to GEO (.25sec) was too long for voice communications, and that the difficulty in developing a working GEO spacecraft was too great. While NASA focused on the LEO approach, Hughes Aerospace began design of a synchronous communications satellite on a spin stabilized bus. Their design became SYNCOM, or synchronous orbit communications satellite, and was launched in July 1963, 18 years after Clarke postulated the idea. The LEO approach to communications was largely abandoned between 1963 and the launch of Iridium in 1997 (Elbert 2008).

After the successful demonstration of the technology by Hughes and SYNCOM, the most pressing requirement that GEO communications satellites were meant to fill was the transoceanic telephone links. An inter-governmental organization (IGO) was formed called the International Telecommunications Satellite Consortium (INTELSAT) to expand upon the burgeoning GEO communications network. INTELSAT evolved to become more of a business than an IGO and numerous satellites were launched to support intercontinental communications (Elbert 2008).

The early satellite communications networks were point to point in nature. They served to provide a long haul connection between two separate networks such as the U.S. and UK telephone networks. They served their purpose in much the same way point to point telegraph connected geographically disparate networks one hundred years before. These early GEO networks were incapable of provisioning services within a single network or country until the 1970s, and they were unable to service any but the largest enterprises (Elbert 2008).

The transition from intercontinental to regional coverage took place in the 1970's. The first country to launch a regional communications spacecraft was Canada, with a single spot beam satellite called ANIK1. The United States followed with a domestic commercial satellite operated by Western Union Telegraph Company called WESTAR in 1974. That same year, NASA demonstrated the first satellite crosslink relay on the ATS-6 satellite. The combination of regional and intercontinental satellite relays and the proliferation of communications satellites allowed for the gradual evolution of a space based circuit switched network that followed the same structure as the telephone networks that it was designed to support (Elbert 2008).

2. Military Satellite Communications

During the same period, as INTELSAT continued to be commercially focused, the Department of Defense recognized that the U.S. military had special communications systems requirements. The goal of these systems was to provide communications between, and to supply information to, military units in situations where terrestrial means of communication are impossible, unreliable, or unavailable while maintaining the ability to re-provision services to support operations. The military satellite communications (MILSATCOM) architecture envisioned consisted of three main programs designed to address three classes of service, wideband, tactical, and protected. Each system evolved as an independent network with different ground terminals and applications (Burch 2011).

The intended customer of wideband satellite services was fixed ground stations or naval ships with large aperture dishes and data rate requirements that are mid to high capacity. The systems that support the wideband mission are Defense Satellite Communication System (DSCS) and Global Broadcast System (GBS), which was replaced by Wideband Global System (WGS). These systems operate in X and Ka bands.

The tactical satellite communications system, sometimes referred to as narrowband, is designed for small aperture, low gain, low to mid data rate users on aircraft, ships and vehicles. The systems include fleet satellite communications system (FLTSATCOM), UHF, UHF Follow on (UFO) and most recently the Mobile User Objective System (MUOS)

Protected MILSATCOM is intended to support mobile users with very small gain antennas with low data rates and a high confidence of connectivity. EHF is not only inherently difficult to intercept or jam, but the protected EHF system is also designed to be survivable in any military environment, including nuclear detonation. EHF and Advanced EHF (AEHF) are designed to provide guaranteed low data rate communications. AEHF also is the only MILSATCOM system to include inter-satellite cross links.

3. Networks

In a fashion similar to terrestrial telephony network providers, the MILSATCOM programs were designed to work independently of each other but with cross connection capabilities at central processing facilities. This methodology has worked well, but is limited in its flexibility to connect users on different systems.

The 1990s saw a paradigm shift in satellite communications. As cellular phones began to proliferate in the terrestrial networks, telecommunications providers began to provide personal services over satellite. Satellite communications became more than just a mechanism for connecting geographically disparate telephone networks, but an end to end communications delivery system unto itself. In the 1990s, the telecommunications world was abuzz with ideas for global satellite phone systems and ideas that included Teledesic's ambitious plan to bring Internet to the world via 840 LEO satellites, and Iridium's global satellite phone coverage constellation of 77 satellites. After the initial commercial optimism that gave birth to such ideas, economic realities caused many plans to be scaled back or canceled all together. Iridium is the

main success story from the 1990s and is an operational satellite phone provider now with plans to replace the current constellation with a new Iridium 2 constellation (Elbert 2008).

From a perspective of networking, there are two fundamentally different architectures around which a satellite network may be built. These two architectures are, a ground based network, and a space based network. As explained by Lloyd Wood in his 2006 paper entitled "Adopting Internet Standards for Orbital Use," the basic difference between ground based and space based networks is where the network layer functionality is performed. In the ground based network, the network functionality is entirely handled by ground stations. In the space based network, the network functionality is handled onboard each satellite.

a. Ground-based networks

In a ground based network, each satellite is simply a re-transmitter where transmission topology is apportioned ahead of time and connections are static. This network can be either a "bent-pipe" where signals are frequency shifted and retransmitted automatically or using signal regeneration and digital processing to cross band platforms or waveforms. Received signals are simply sent straight back to another location on the ground. This allows ground stations and users to exchange information with each other when they are in the same footprint of the satellite. In order to reach users outside of that geographic area, terrestrial networks must provide transport. In this sense, a ground based system provides a "last hop" for an existing terrestrial network. This causes challenges in the space segment for media access control (MAC), and logical link control (LLC) sub layers of the data link layer and as a result, handover issues can be a problem for any multipoint network running multiple network protocols (L. Wood et al. 2006).

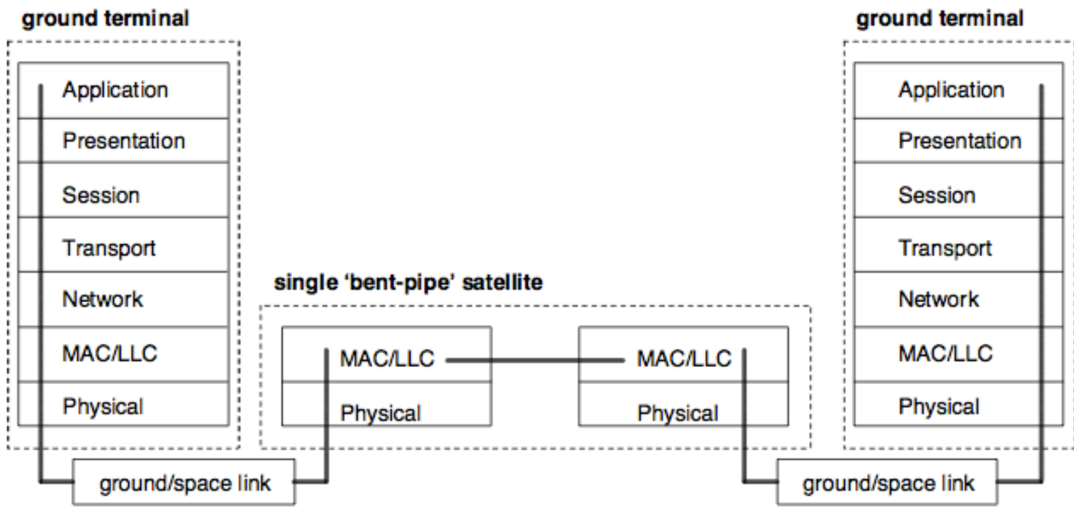

Figure 2. Open Systems Interconnect (OSI) Model (from L. Wood et al. 2006)

Since the satellites are only enabling the last hop of the network, the network topology on the ground is entirely arbitrary and likely to be influenced by factors other than the satellites such as political, geographic or economic factors. It can be safely assumed that in a ground based network, all of the TT&C ground stations will be networked so that they may share information about the health and status of the constellation at any moment. Beyond this, there are a large number of possible network topologies that could connect each gateway with existing networks, including the Internet. As a result, the ground based satellite network topology is governed by the considerations of the terrestrial network, whereas the satellite based network topology is heavily dependent on orbital geometry (L. Wood et al. 2006).

b. Space-based Networks

In space based networks, according to Wood, each satellite in the constellation has a router onboard to perform processing and routing to communicate with neighboring satellites as well as ground stations. The communications with neighboring satellites is accomplished using inter-satellite links(ISL). This type of network allows a user on the ground to communicate with

a ground station beyond the line of sight of the satellite receiving the uplink, or with users below distant satellites without requiring local gateways below each satellite or large footprint terrestrial infrastructures.

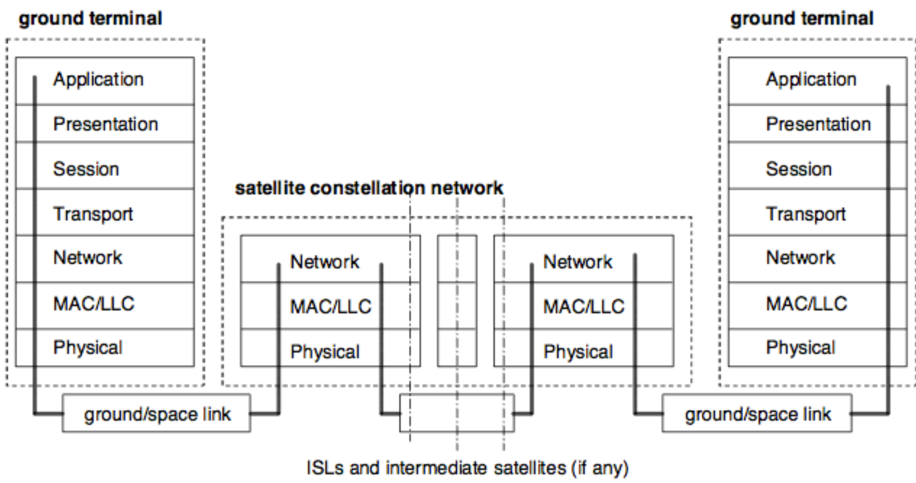

Figure 3. OSI Model Satellite Specific (from L. Wood et al. 2006)

By using ISLs, the space segment reaches the network layer of the OSI model allowing for the consideration of mesh communication between a constellation of satellites in orbit without central gateway ground stations providing the processing. Such satellites will require both routing and switching and in this case the constellation is itself a full network and by including the associated ground stations, is an autonomous system. By contrast, in the ground based system, each gateway could be considered a distinct autonomous system (L. Wood et al. 2006).

For any constellation utilizing circular orbits, it is possible to use fixed links between satellites ahead and behind to satellites in stationary relative positions for intra-planar communication. The same fixed links cannot be used for inter-planar links between satellites in different orbital planes as the relative position and velocity differ as the satellites converge and diverge at crossings

points, typically over the poles. The changing relative position also inputs Doppler shift into the link requiring specialized radios and receivers to adapt to the changing frequencies.

c. Comparison of ground-based and space-based

The space-based approach utilizing ISLs decreases ground network traffic as well as the need for multiple ground-space hops across limited links. However, the processing required to route and switch from the satellite is significantly more complex than the processing from a ground station. By placing much more complex processing systems on orbit, the space segment can become much more automated, allowing more enterprise service to users despite their geographic location or proximity to existing gateway infrastructure.

Conversely, the ground-based network, utilizing bent-pipe links to and from satellites is limited to operating in areas where users and gateways are concurrently in contact with the satellite (L. Wood et al. 2006).

The space-based network is constrained by the orbital geometry of the constellation, as well as the difficulties of providing the required processing to adapt to the constantly changing topography. In addition, by pushing network functionality onto the satellites, issues in the network layer and in the space segment must be considered concurrently (L. Wood et al. 2006).

The ground based network separates network functionality from the space segment, allowing issues in the space segment and the network layer to be considered separately (L. Wood et al. 2006).

d. TCP/IP Networking in the Space Segment

To date, all satellite networks have either used ground based architectures, such as MUOS, MILSTAR, INTELSAT, Inmarsat, etc., or specialized protocols designed specifically for a certain constellation, such as the Iridium constellation. There have only been a handful of attempts to use standardized TCP/IP networking protocols on satellite nodes.

The first use of an TCP/IP router onboard a satellite was part of a demonstration of the Office of the Secretary of Defense (OSD) space-based network centric concepts and major elements of the National Reconnaissance Office (NRO) and Transformational Communications Architecture (TCA) using technology based on Cisco Internet routers. The demonstration also illustrated that the functional intent of the Consultative Committee for Space Data Systems (CCSDS) Space Link Extension (SLE) was met. The key point of the demonstration was to show the ability to securely use networks to control infrastructure owned or controlled by various parties (Ivancic, et al. 2005).

On September 27th, 2003, a Cisco Internet router was launched into LEO onboard the United Kingdom Disaster Monitoring Constellation (UK-DMC), a small satellite built by Surrey Satellite Technology Limited (SSTL) which carried various mission payloads including remote sensing imaging sensors, an experiment demonstrating GNSS reflectometry, and a water resistojet propulsion system. The demonstration was called CLEO for Cisco Router in Low Earth Orbit. The router was tested and shown to allow communications with different payloads in space (Ivancic, et al. 2005).

While the UK-DMC satellite's main purpose was to provide images of the environment on Earth, the secondary payload, CLEO, was a primary focus of experimentation by a wide range of organizations including Cisco Systems, SSTL, the U.S. National Aeronautics and Space Administration (NASA), the U.S. Department of Defense (DoD) including each of the four military branches, General Dynamics advanced Information Systems, Universal Space Network Inc, Western DataCom, and others. The router was used as an IP compliant space based asset for the OSD rapid Acquisition Netcentric Virtual Mission Operations Center (VMOC) demonstration as well. IP based software was used to acquire satellite telemetry, request images from SSTL's satellites dynamically, and perform real time access to onboard equipment. The VMOC test allowed ground users in the field to dynamically task and receive imagery from the satellite in June of 2004 (Ivancic, et al. 2005).

The essence of the CLEO experiment was to use a router and IP software to turn the various subsystems and components on the satellite into a small private network, accessible via standard IP protocol traffic. This allows standard payloads to be placed on the onboard network and commanded via standard commercial IP from multiple dislocated ground sites and systems (Ivancic, et al. 2005).

The second use of a TCP/IP capable spacecraft was the Internet Router In Space (IRIS) payload launched on INTELSAT14 in 2009. IRIS consists of a Cisco router, designed and built for the space environment and intended to provided IP routed communications for communications satellite customers using the C and Ku band payloads on INTELSAT14. The performance of the IRIS payload was demonstrated during a Joint Capabilities Technology Demonstration (JCTD) in 2010. The Operational Utility Assessment (OUA)of the IRIS JCTD provided significant insight into the capabilities that IRIS may provide to the military user and concluded that the IRIS payload provided useful capabilities (Cuevas, et al. 2010).

The JCTD was analyzed by the Johns Hopkins University Applied Physics Laboratory which had the following to say about the impact of IRIS:

> IRIS provided direct connectivity to different user groups, terminal types, and user nodes located in distinct geographical areas. Cross-band, inter-beam, and intra-beam connectivity was seamlessly achieved once the user nodes and satellite ground terminals were properly configured. The network supported Internet access as well as VPN connectivity.
>
> The end-to-end QoS [Quality of Service] architecture implemented differentiated services and effectively demonstrated interoperability of QoS capabilities. The application performance results showed that the QoS features provided a tiered service class structure based on individual marking of packets and the ability to better control performance experienced by the end-user. The quality of VoIP [Voice Over IP] and VTC [Video Teleconference] was acceptable and remained consistent, even under heavy traffic conditions. On the other hand, the performance of FTP, web applications, and chat, marked as BE varied widely, depending on

> the traffic loading conditions. However, as demonstrated in OD-4, the ability to utilize the QoS features with some common real-world applications was found to be difficult and problematic. These applications did not natively support differentiated services as a feature or did not support them very well.
>
> The BoD [Bandwidth on Demand] capability allocated resources dynamically, upon request, to all terminal nodes. However, the response to sudden temporary surges of traffic was not always satisfactory, especially for real-time VTC. The amount of overhead added by the LM is high (close to 60%) and it significantly affects the overall capacity.
>
> Overall network performance was enhanced by the use of Cisco WAAS [Wide Area Applications Services] units installed on each CGR[Cisco Ground Router]. The units improved the performance of TCP traffic, particularly for large size FTP [File Transfer Protocol] files by providing protocol enhancement, compression, and caching. (Cuevas, et al. 2010).

Both the CLEO and IRIS demonstrations indicate that there are both significant benefits and significant problems associated with adoption of Internet Protocols to satellite communications. The advantages are many and are mirrored in the long evolution of terrestrial communications; flexibility, uniformity, interoperability, and scalability are all hallmarks of networked communications. The problems are inherent in the behavior of satellite nodes with regards to network topology. Orbital nodes have a unique mixture of static and dynamic properties and the dynamic properties are often highly cyclical and thus predictable, which do not lend themselves well to any existing routing protocol. While IRIS and CLEO demonstrate the many benefits, it is only the first step. A single networked satellite node will behave differently from a constellation of satellites, networked via inter-satellite links.

THIS PAGE INTENTIONALLY LEFT BLANK

III. CLASSICAL ORBITAL ELEMENTS AND SPACECRAFT GEOMETRY

In general, orbits can be broken down into four different class based mainly on the altitudes in which they cover, with the exception of the Highly Elliptical Orbit (HEO). Each class of orbit has its own strength and weakness that will be examined in detail later in this chapter. LEO exists from roughly 160 km to 2000 km (Curtis 2009). Medium Earth Orbit (MEO) is defined as the range above 2000 km and below 30,000 km. Geosynchronous Orbits (GEO) are defined as those circular orbits whose period is equal to one sidereal day, an altitude of 35,788km. A Geostationary orbit is a special case of a geosynchronous orbit with an inclination near zero. HEO orbits are those orbits that have a low perigee and a high apogee. Generally the amount of power required from a ground terminal with a given data rate and modulation scheme is much less to reach a satellite in LEO than GEO. This is because the radiated power falls off at the rate of 1/R^2 where R is the distance between the two antennas.

A. CLASSICAL ORBITAL ELEMENTS

Six elements are used to describe an orbit, referred to as the Classical Orbital Elements (COE). A general understanding of the COE will be useful to describe the process for orbit selection and performance evaluation. The terms and their definition are provided in Table 1. Additionally they are depicted in Figure 4 to help visualize them.

1. Semi major Axis (a)

Semi major axis defines the size of the orbit. It is the distance from the center of the earth to apogee plus the distance from the center of the earth to perigee divided by two (Wertz and Larson 2007).

$$a = \frac{r_a + r_p}{2} \tag{0.1}$$

2. Eccentricity (e)

Eccentricity describes the shape of the ellipse. It is essentially a ratio of the sum and the difference or radius of apogee and the radius of perigee (Wertz, Everet and Puschell 2011).

$$e = \frac{r_a - r_p}{r_a + r_p} \tag{2.2}$$

3. Inclination (i)

Inclination is used determine the tilt of the orbit. It is the angle between the equatorial plane and the orbital plane. Equatorial orbits have inclinations equal to 0° or 180°. Polar orbits, those orbits that cross both poles with each orbit, have inclinations of 90°. Prograde orbits have an inclination of 0°≤ i ≤ 90°. Most orbits are prograde because it is cheaper to launch satellites in these orbits. Retrograde orbits have inclinations of 90° ≤ i ≤180°. The benefit of a retrograde orbit is that the satellite is moving in the opposite direction of the rotation of the earth and can therefore reduce the time it takes for a satellite to pass over the same target on the surface of the Earth. This is called revisit time (Wertz, Everet and Puschell 2011).

4. Right Ascension of Ascending Node (Ω)

The right ascension of the ascending node is point along the equator where the satellite crosses from the southern hemisphere to the northern hemisphere. It is the angle from the vernal equinox to the ascending node (Wertz, Everet and Puschell 2011).

5. Argument of Perigee (ω)

The argument of perigee is the angle from the ascending node to perigee. The argument of perigee is used to determine the orbits orientation within the orbital plane (Wertz, Everet and Puschell 2011).

6. True Anomaly (v)

The true anomaly is used to determine the spacecraft's location within the orbit. It is the angle from perigee to the spacecraft's location (Wertz, Everet and Puschell 2011).

Element	Name	Description	Range	Undefined
a	Semimajor axis	Size	Minimum of 6508	Never
e	Eccentricity	Shape	0<e<1	Never
Ω	Right Ascension of ascending node	Angle from vernal equinox to ascending node	0≤Ω≤360°	When i=0 or 180° (equatorial orbit)
i	Inclination	Tilt of orbital plane with respect to equator	0≤i≤180°	Never
ω	Argument of Perigee	Angle from ascending node to perigee	0≤ω≤360°	Never
v	True Anomaly	Angle from perigee to spacecraft's position	0≤v≤360°	Never
r_p	Radius of Perigee	Distance from center of earth of perigee		Never
r_a	Radius of Apogee	Distance from Center of earth to apogee		Never
P	Period	Time require to complete one orbit		Never

Table 1. Classical Orbital Elements

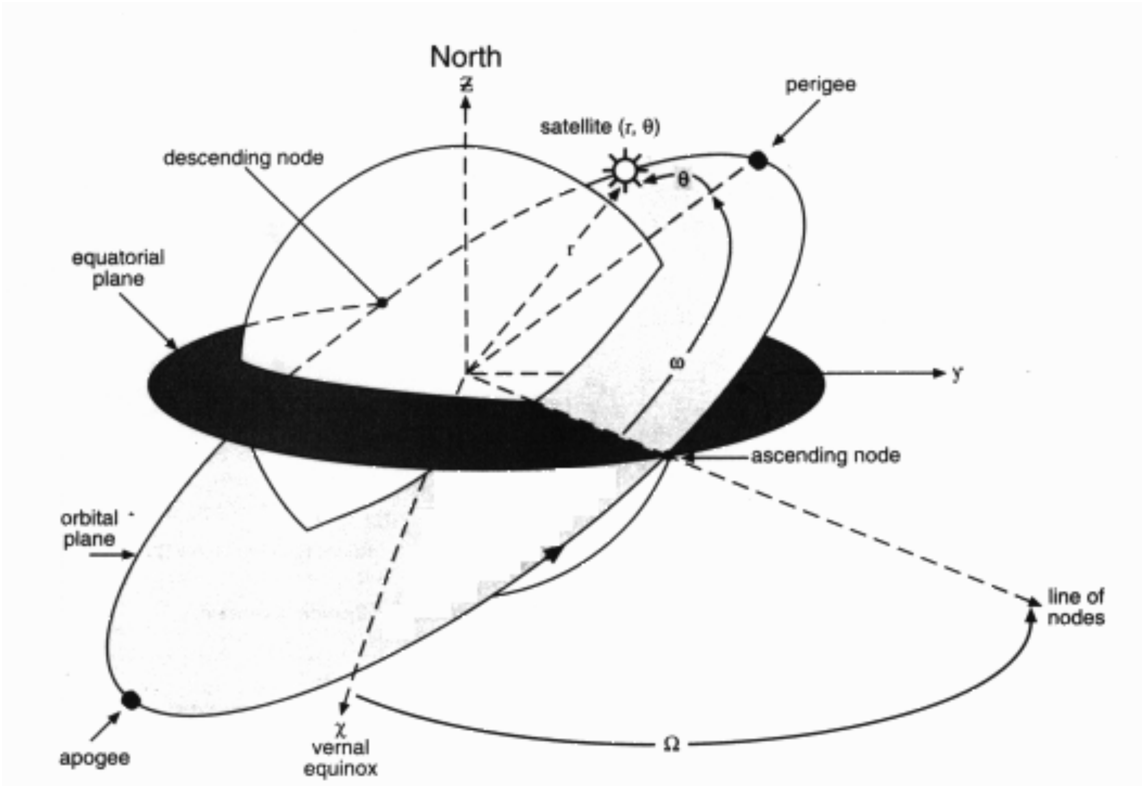

Figure 4. Classical Orbital Elements (from Hess 2012)

Perigee is defined as the point in the orbit where the satellite is closes to the earth while apogee is when the satellite is farthest from the earth. In circular orbits the radius of perigee and the radius of apogee will be equal to one another.

B. ORBITAL TYPES

It is important to identify trade spaces in an effort to engineer several baseline solutions so they can be compared to one another in order to identify the best solution from a cost-performance stand point.

1. Low Earth Orbit (LEO)

NASA defines Low Earth Orbit (LEO) as those orbits that range from roughly 150km to 2000 km. Satellites in this orbit enjoy the benefit of a lower power requirement for transmissions to and from the ground. Additionally, remote sensing satellites in this orbit have a better resolution. Satellites are cheaper to

launch into this orbit. However, because of factors such as drag, the Van Allen Radiation Belts, and atomic oxygen, satellites in LEO have a shorter lifetime than spacecraft operating in higher orbits. As a result of their altitude, satellites in this orbit have a smaller field of view and coverage areas. This must be taken into consideration when designing, planning, and launching a constellation of satellites. Meaning the satellites must be built and launched in relatively quick succession to ensure the entire constellation is built out and the final spacecraft is online and operational before the first satellite launched is reaching the end of its design life (Wertz and Larson 2007).

2. Medium Earth Orbit (MEO)

The Medium Earth Orbit includes all orbits from 2000km to 30,000 km. Typically they reside around 20,000 km. Because of their altitude satellites in this orbit have a larger field of view and coverage area, and therefore can require fewer satellites to provide complete global coverage. Conversely the require more power to transmit to and from the ground, and have a lower resolution the LEO remote sensing satellites.

3. Geosynchronous Orbit (GEO)

The geosynchronous orbits have an altitude of around 35,678 km. The benefit of this orbit is that the orbital period, the time it takes the satellite to rotate around the earth, is the same time it take takes the earth to rotate on its axis which is one sidereal day or 23.94 hours. This means the satellite is constantly "staring" at the same point on the earth's surface, which is a huge benefit for both communications and remote sensing satellites. It also means a GEO satellite can see roughly one third of the earth at one time, resulting in fewer satellites to provide complete global coverage. However, it also means that the resolution for remote sensing satellites is much lower at GEO than at lower altitudes (given the same optics). Likewise a significantly more amount of power is required to transmit between the satellite and the ground, which translates to a larger

spacecraft. GEO satellites are very expensive to launch both because of their size and the distance they have to travel to get into their operational orbit.

4. Highly Elliptical Orbit (HEO)

A highly elliptical orbit is an orbit that has relatively low perigee and a relatively high apogee resulting in a large eccentricity value (approaching 1) and can pass through all other orbital altitudes. There is one special case of HEO called a Molniya orbit that has a 12 hour period, an eccentricity value of approximately 0.75 and inclinations of 63.4° or 116.6°. The benefit of the Molniya orbit is that nearly 11 hours of its period is spent at apogee which proves to be very useful in providing coverage for higher altitudes (Wertz and Larson 2007).

5. Parking Orbit

Satellites occupy parking orbits only for a short while. A parking orbit is the equivalent of a holding pattern for an aircraft. The satellite remains in the parking orbit until the optimal conditions occur to enter the transfer orbit. Often, some spacecraft subsystems are turned on to ensure the craft survived the launch phase. However, because the spacecraft will be executing another burn to reach its operational orbit, solar arrays and antenna remain stowed to prevent damage.

6. Transfer Orbit

A transfer orbit is simply an elliptical orbit a spacecraft is in between the parking orbit and its operational orbit.

C. CURRENT SATELLITE COMMUNICATIONS CONSTELLATIONS

The purpose of this part of the discussion is to introduce and discuss the constellations that are currently in use by satellite communications companies that will serve as both the foundation for the constellation discussion and highlight areas that can be optimized for the particular problem set this project is attempting to solve.

1. Iridium

Iridium is a constellation of communication satellites in Low Earth Orbit that provides global coverage telephone services to its users. It consists of 66 satellites in 6 orbital planes in a polar orbit. There are eleven satellites per plane. The planes are spaced out 31.6 degrees apart with planes one and six being 22 degrees apart. Each satellite in planes two through five maintains four Inter-Satellite Links (ISLs), one forward and aft in the same plane and one on either side to satellites in adjacent planes. Because satellites in planes one and six are traveling in opposite directions they do not maintain ISLs between planes one and six.

The constellation is due to be replaced beginning in 2015. Essentially, Iridium Next will completely replace the current iridium constellation in the roughly the same orbits. The altitude of the current Iridium constellation is roughly 770km with the new constellation to occupy an altitude of 780 km. Unlike most communications satellites, Iridium satellites chose to put the switching technology onboard the satellite. Had the designers not done so there would have been two choices, put a ground station that contains the switching and routing technology in each "cell" of coverage throughout the entire globe, or several ground stations throughout the globe to route calls. Obviously there are ground stations incorporated into the Iridium design that allow the satellites to interface with the terrestrial telephone and data services. The benefit of having switching technology onboard the satellite is that the satellites can determine the best path to the closest ground station resulting in much shorter delays and a faster network.

Iridium Next is offering hosted payloads to help offset the cost the cost of the constellation. Each satellite will have excess size, weight, and power that will be available to host additional payloads. Currently the projected SWAP that will be available is 50 kg dimensions of 30 x 40 x 70 cm, 50 W average power 200 W peak power and will have a data rate of up 1 Mbps (Iridium 2012). The current iteration of the Internet Router In Space (IRIS) is too large to be a hosted payload

on Iridium Next. For the purpose of this project the assumption will be made that the hardware for an Iridium type IP switched network will exist in the future.

2. Inmarsat

Inmarsat is a company that provides worldwide satellite coverage enabling both voice and data services to its clients. It is important to clarify the distinction between global and worldwide. Inmarsat's satellites reside in the Geo belt. Global coverage means that a customer, regardless of where they are on globe has access to the services a given network provides. Worldwide coverage generally means coverage is available between the latitudes of 70 N and 70 S. It is between these latitudes where most of the world's population lives therefore many commercial providers opt for worldwide vice global coverage. Currently Inmarsat is operating 11 satellites in the geosynchronous belt. Eight of these satellites are what Inmarsat refers their I-3 spacecraft. The remaining three satellites are called the I-4 and represent a significant improvement over the rest of the aging constellation. The I-4 Satellites have a nine meter parabolic antenna and 120 helix element combined into a single array. The onboard digital processor controls the antennae, beam forming and channel allocation. Inmarsat's website advertises voice and circuit switched data services. Because their satellites are located in the geosynchronous belt the terminal devices that interface with the network are much larger than Iridium. Again this is because radiated power drops off at the rate of 1/r^2 where r is the distance between the two antennae. Inmarsat is addressing this issue with their I-5 satellites. The I-5 satellites will operate 89 Ka spot beams and offer data rates of up 50 mbps, and antennae as small as 60 cm (Inmarsat 2012).

3. Terrestar

Terrastar is another telecommunications company that offers satellite base phone service. While their use of the geosynchronous belt is not uncommon, their satellite does have a rather unique feature. It has an 18-meter parabolic S band reflector that enables the terminal devices to be as small as a Blackberry (TerreStar 2012).

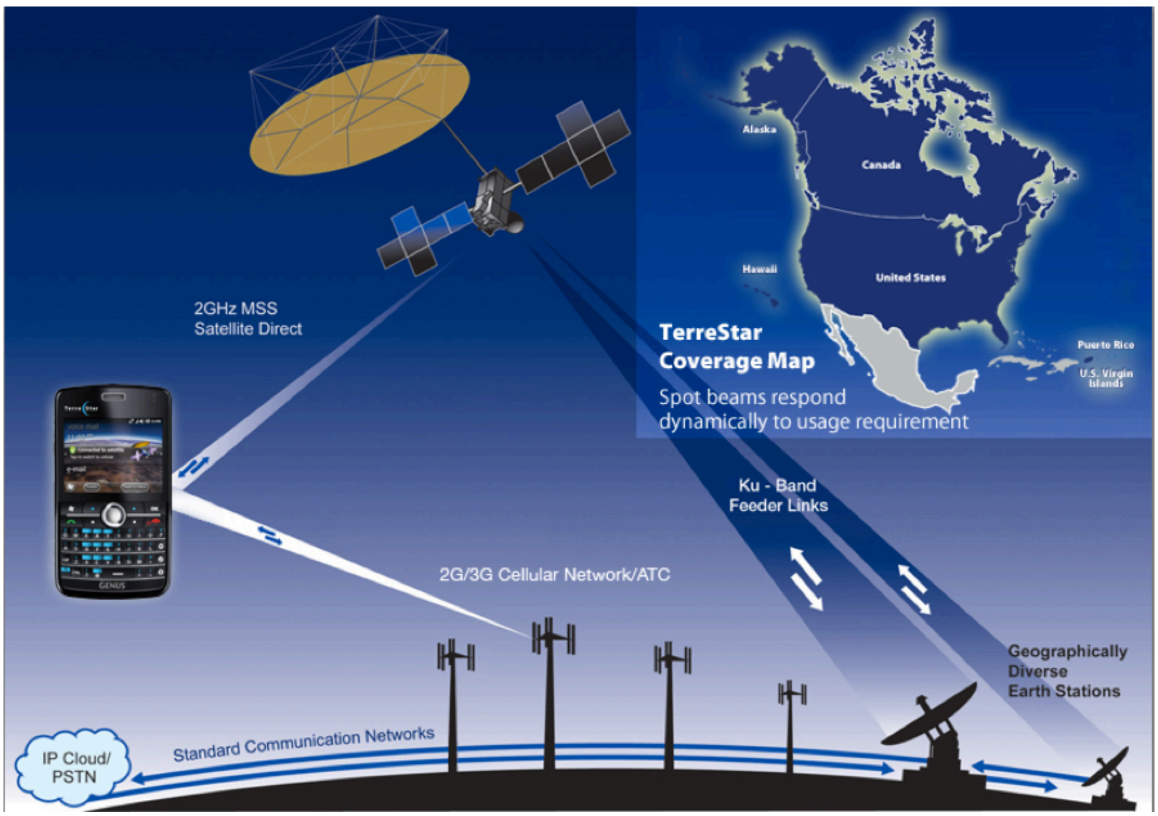

Figure 5. Terrastar OV1

D. SPACECRAFT GEOMETRY

The architectures being conceived, modeled, and analyzed are intended to support the warfighter in an attempt to meet the ever-growing appetite for data and communications on the move. Having said that, and in an effort to reduce the problem set, the terminal devices the architecture will be intended to interface with will be restricted to those already in the inventory of the U.S. Military; in

particular to those devices that can or would be able to interface the Navy's Mobile User Objective System (MUOS). It is these radios and antennae that will be used to design the constellations. The orbital regimes that will be examined will be restricted to low earth orbit (LEO) and geosynchronous (GEO). Again because the architecture will be designed to support the warfighter it is a requirement that the architecture provide worldwide coverage.

Because it is important to ensure a common language as the discussion moves forward it is import to define a few terms. This section will introduce and define the terms commonly known as Spacecraft Geometry, which will be extremely useful when discussing the design of the respective architectures.

1. Nadir Angle: η

The Nadir angle, η is measured at the spacecraft from sub satellite point to the target (Wertz and Larson 2007). The target can be an object to be imaged or a receiver that the spacecraft is trying to communicate with.

2. Angular Radius of The Earth: ρ

The angular radius of the Earth is measured at the spacecraft from the sub-satellite point to the true outer horizon of what the spacecraft can effectively "see."

3. Maximum Earth Central Angle: λ_0

The earth central angle is measured from the center of the earth to the true outer horizon of what the satellite can "see."

4. Elevation Angle: ε

The elevation angle is measured at the target and is the angle between the spacecraft and the local horizon.

5. Altitude: H

The altitude is the height of the satellite measured from the sub-satellite point to the satellite.

6. Slant Range: D

The slant range is the range from the satellite to the target. Figure 6 provides a graphical representation of the elements related to spacecraft geometry.

7. Earth Central Angle λ

λ is the angle measured at the center of the earth from the earth's center to the target.

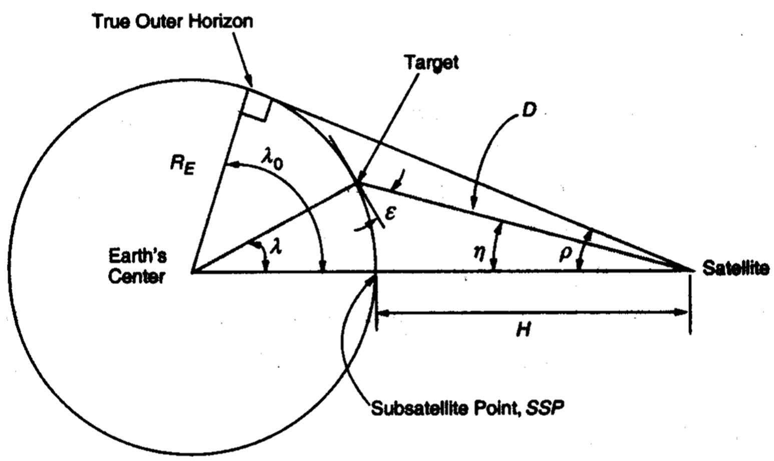

Figure 6. Spacecraft Geometry (from Wertz and Larson 2007)

8. Access Area

The term access area is used to describe the total area on the ground that a space instrument could potentially see at a given time.

9. Footprint

The terms footprint and Field of View (FOV) are often used interchangeably. Both are used to describe the area on the ground that a spacecraft instrument is actually able to see.

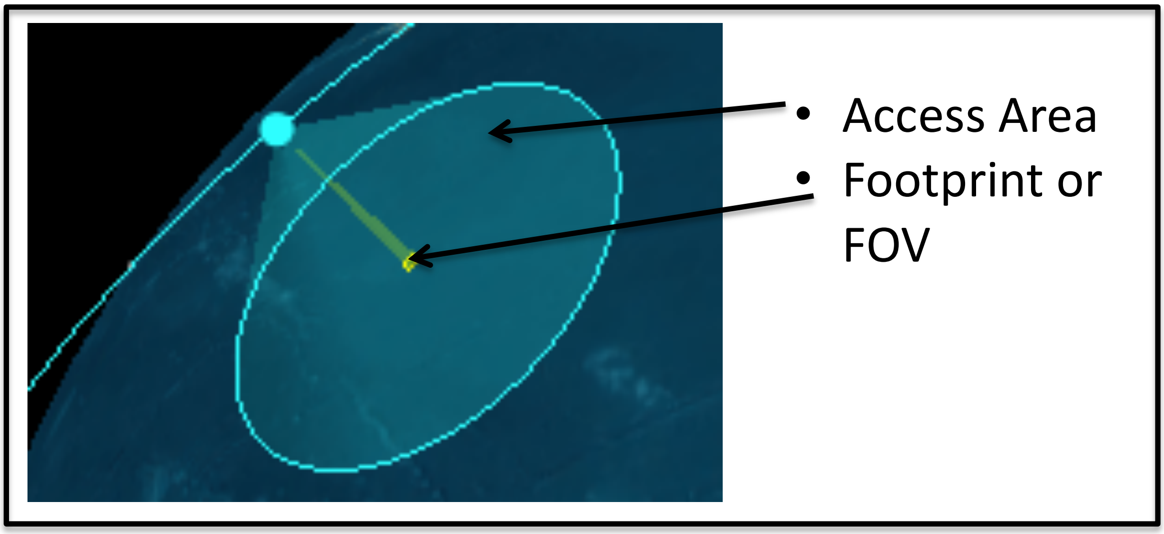

Figure 7. Access Area and Field of View

E. CONSTELLATION DESIGN

Constellation Design is a very complicated process. The mission the system is trying to accomplish is what drives the design process of not only the spacecraft but the orbit(s) and constellation as well. When launching a constellation of satellites or even a single satellite, cost is a constraint that must be considered and minimized. Because the main focus of this project is to develop constellations of satellites that are capable of providing global coverage, launching a minimum of three satellites in geosynchronous orbits will fulfill this need. However, it may not be the best answer. As stated in the original problem statement, this project is focusing on delivering high bandwidth, IP routed communications to the disadvantaged user, and users on the move while minimizing the logistical footprint of the equipment. In other words, the goal is to

push data from a satellite to handheld or man portable unit. That being said, the governing equation driving the design of the constellations is the link equation.

1. Access Area and Number of Satellites.

Much thought was given to the design of the low earth orbit constellation. The primary benefits of using a low earth orbit are that the end user terminal devices can be much smaller and are required to transmit much less power in order to close the link with in a specified set of parameters. This benefit is not without cost, however. The closer the satellite is to the earth, the smaller the coverage area of the satellite is and therefore the more satellites that are required to have 100% global coverage. The coverage area for a satellite pointed directly at earth is calculated by the equation below.

$$A = 2\pi R_e^{\,2}(1-\cos\lambda_0) \tag{0.2}$$

Where A is the area in square meters, R_e is the radius of the earth and λ_0 is the maximum earth central angle. From the equation below it is evident that the maximum earth central angle increases as the altitude of the satellite increases.

$$\cos^{-1}\left(\frac{R_e}{R_e+h}\right) \tag{0.3}$$

As a baseline calculation, a minimum number of satellites required to obtain an aggregate surface area equal to that of the earth can be obtained by simply dividing the surface area of the earth by the area the satellite can see. Notice this does not mean that 100% global coverage is achieved by using this number of satellites as it does not take into account the spherical shape of the coverage. Nor does it take into account the amount of overlap in coverage areas required between satellites in order to ensure the handoff of a connection from a ground user from one satellite to the other is executed. This simply serves as sanity check. The coverage area of a satellite as a function of altitude is capture in Figure 8 Because there will need to be some amount of overlap between

satellite coverage areas to ensure persistent global coverage a quick calculation can be performed to get the number of satellites required. For example if you determine that 50% of a single satellite's coverage are is required to be seen by at least one other satellite to ensure persistent global coverage simply double the amount found. Likewise, knowing the number of satellites in the Iridium constellation, as well as their operating altitude, it can be calculated that 73% of a single Iridium satellite coverage area is covered by at least one other satellite. The corresponding minimum number of satellites to achieve an aggregate surface area equal to that of the earth is captured in Figure 9.

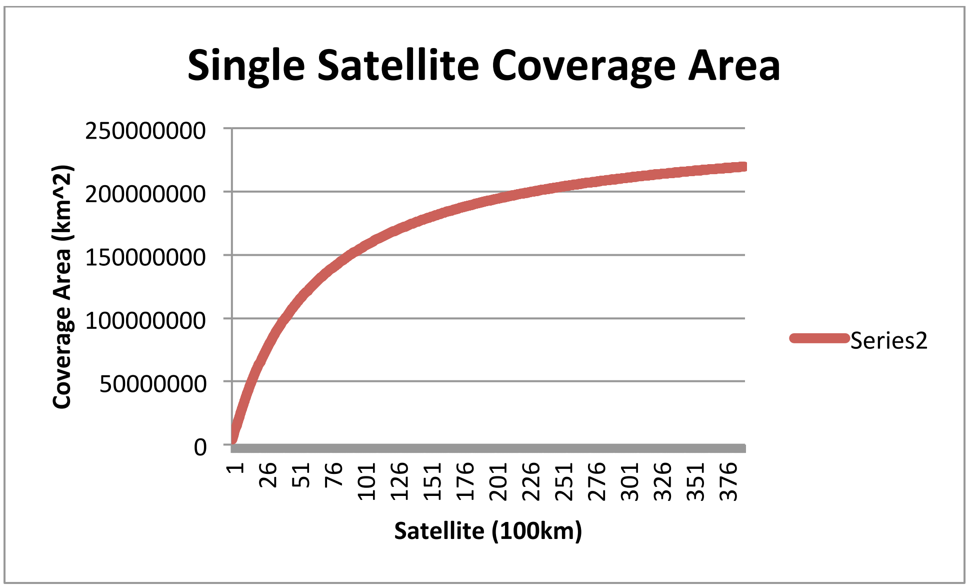

Figure 8. Single Satellite Coverage Area

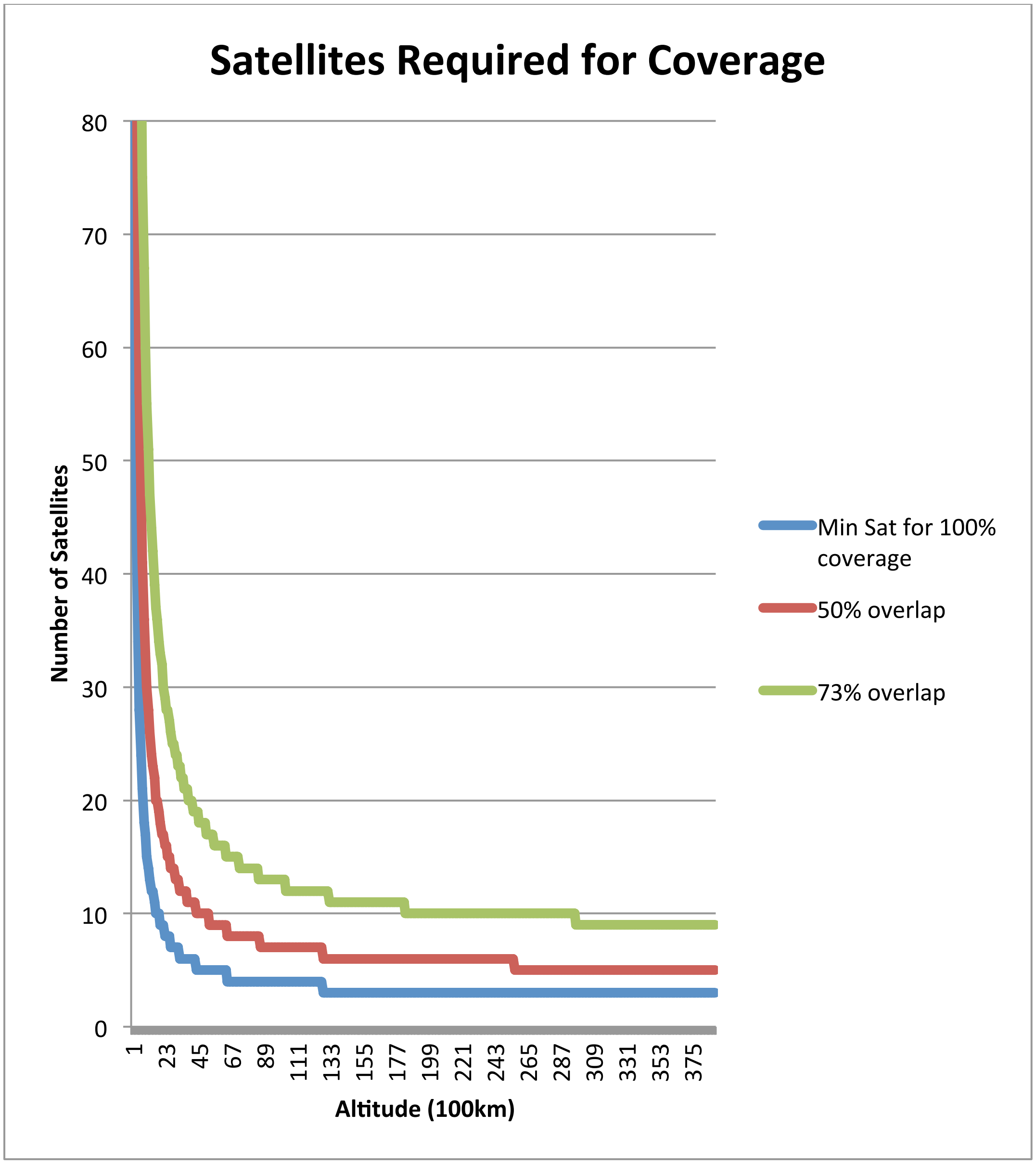

Figure 9. Satellites Required for Coverage

2. Low Earth Orbit Constellation

Now that the access areas and minimum number of satellites have been calculated, the arrangement of the satellites can be determined and discussed. The first approach to solve this problem was to use a walker constellation. *Space Mission Engineering: The New SMAD* by Wertz and Larson describes how walker constellations are designed. Essentially, the total number of satellites is

divided by a given number of planes. Next determine the spacing between adjacent planes. Finally determine the spacing within same plane between satellites. It is important to note that the inclination of all the satellites is the same (Wertz, Everet and Puschell 2011).

T/P/F – Walker Delta Patterns
t= Number of Satellites p= Number of orbit planes evenly spaced in node. f= Relative spacing between Satellites in Adjacent planes. Define s= t/p= Number of satellites per plane Define Pattern Unit, PU = 360/t Planes are spaced at intervals of s PUs Satellites are spaced at intervals of p PUs within each plane If a satellite is at its ascending node, the next most easterly satellite will be f PUs past the node. f is an integer from 0 to (p-1). Example: 15/5/1 15 satellites in 5 planes (t=15, p=5) 3 satellites per plane (s = t/p = 3) PU = 350/t =360/15 = 24 deg. In plane spacing between satellites = PU*p = 24*5 = 120 Node Spacing = PU*s = 24*3 = 72 degrees Phase difference between adjacent planes = PUs*f = 24 *1 = 24 deg

Table 2. Characteristics of a Walker Delta Pattern Constellation (from Wertz, Everet and Puschell 2011)

Typically Walker constellations have similar orbits so that pertubations will affect each satellite and respective plane in a similar fashion. However, the Iridium satellite constellation is not a true Walker constellation, as its planes are not all evenly spaced. Planes 1 through 6 are evenly spaced when moving from plane 1 to 2, 2 to 3, and so on. However, moving from plane there is a greater space moving from plane 6 to 1. The Iridium constellation is an Adams and Rider constellation in which the eccentricities of the orbits are adjusted ever so slightly to ensure the satellites do not drift and or coalesce towards one another (Adams and Rider 1987),

3. Geosynchronous Constellation Considerations.

Due to the altitude the geosynchronous satellites operate, they “see” significantly more of the earth at a time. Figure 10 shows that with 50% overlap 3 satellites are required to provide world wide coverage (70 north to 70 south latitude). Because the geosynchronous orbit slots are such a valuable commodity, they must be controlled in order to prevent countries from launching satellites into orbit and increasing the risk for collisions and creating a debris field that could potentially have a significant impact on neighboring satellites. The International Telecommunications Union (ITU) is the controlling organization with which all satellite owners must request a geosynchronous orbit slot from. To ensure a high fidelity model was being created for the research team, the team used orbital slots that are owned by the United States.

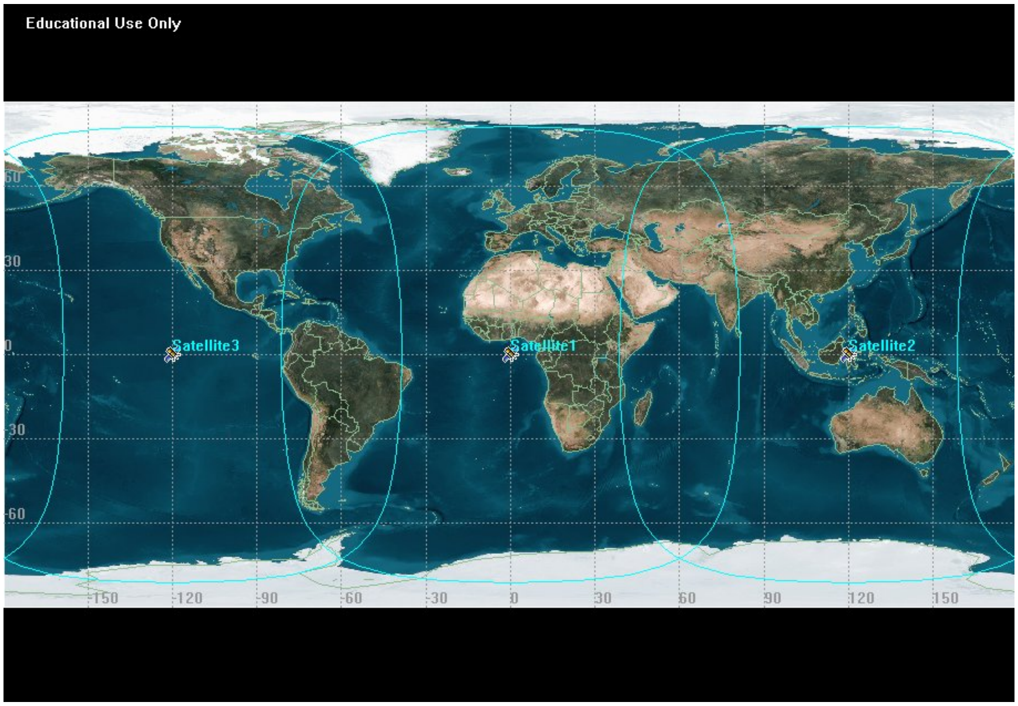

Figure 10. Three Ball Geo Constellation

a. Geosynchronous Constellations

The United States Government has several existing constellations in geosynchronous orbits. Primarily the constellations exist to provide communications for the Department of Defense and the Military. Military communications satellites can be divided up into three main categories; UHF (narrowband communications), SHF, EHF. The latter of the two comprise the wideband portion of the Military satellite communications. These satellites are carefully positioned to ensure the coverage is where it is needed. Ultra High Frequency Follow on (UFO) is currently the primary provider of narrowband communications for the military. The current positions are located in the Figure 11. Wideband communications is comprised of DSCS, and WGS. Figure 12 depicts the approximate positions of the satellites that comprise these respective constellations.

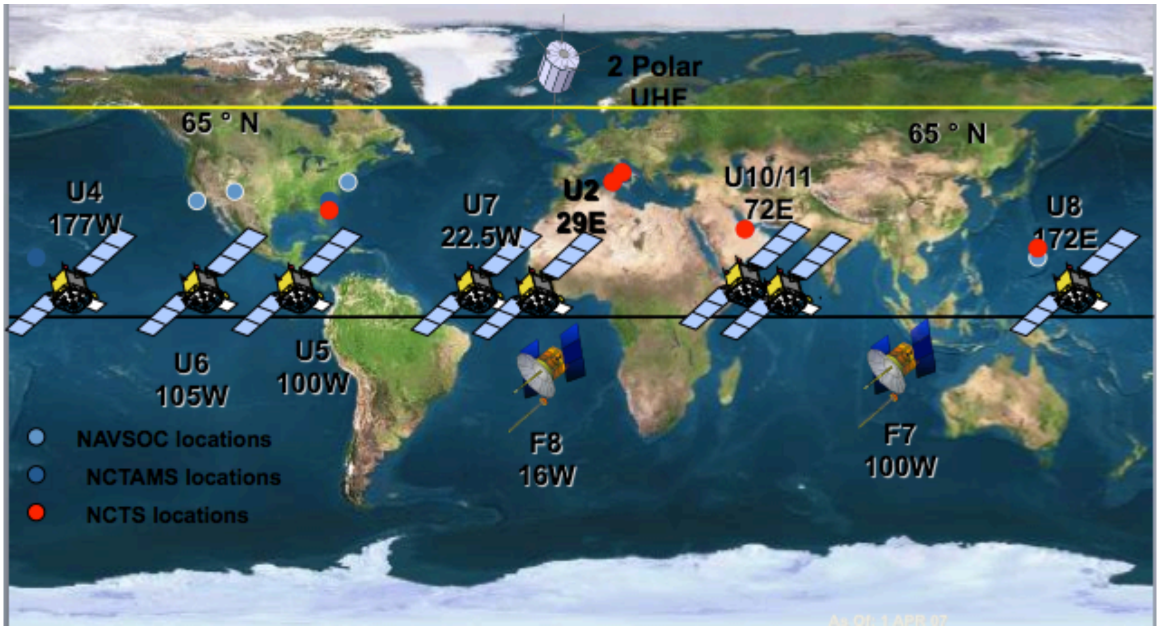

Figure 11. Narrowband Satellites

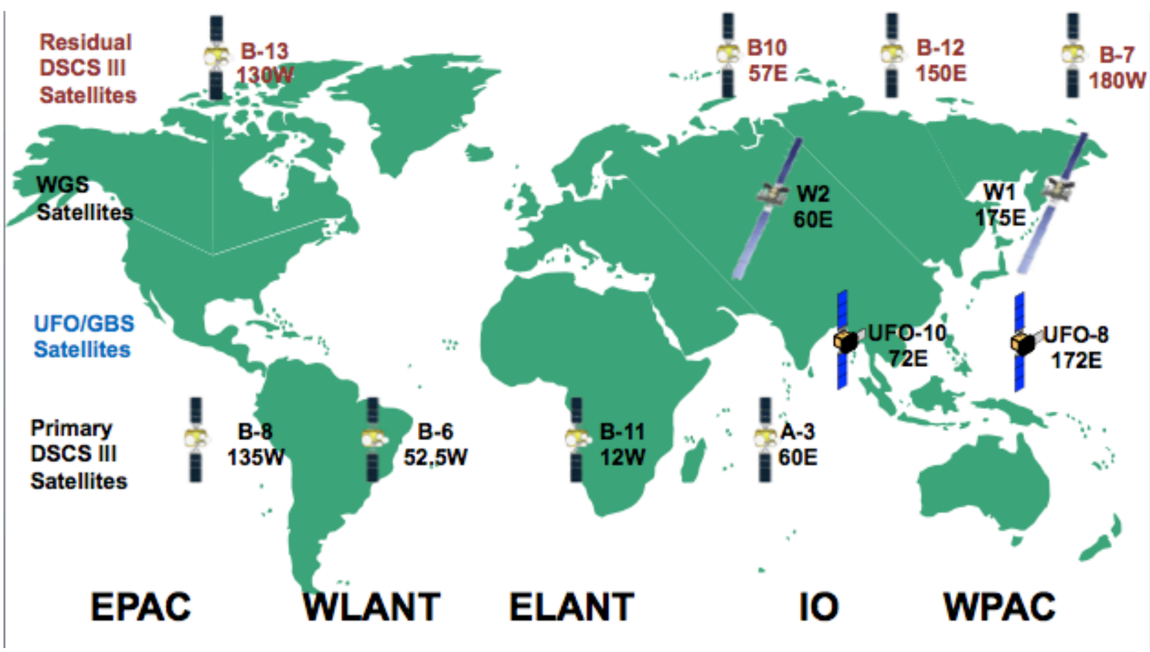

Figure 12. Wideband Satellites

b. *Other Constellations*

There were other constellations that were considered but not used. The GPS constellation offers persistent global coverage, however, due to the fact that it operates in a Medium Earth Orbit, it was determined there was no significant benefit over a Low Earth Orbit Constellation or a constellation operating in geosynchronous orbit. There are a number of commercial communications satellites that were previously mentioned that were not considered for use in military only constellation, however some of the constellations were examined for the purposes of hosted payload. A hosted payload is a payload that is occupied by the excess size weight and power (SWAP) of a host satellite. The owner of the host satellite sells the valuable real estate on the satellite and supplies the hosted payload with power. By utilizing hosted space on satellites that are being launched, communication assets can be launched into orbit at a much-reduced cost. Hosted payloads will be examined in another chapter but it represents a unique opportunity for the Military to save money and still get the valuable communications services it needs.

IV. COMMUNICATIONS LINKS

A. LOW EARTH ORBIT LINK

The link equation describes the conditions in which a particular radio frequency (RF) link operates with the intent to maintain a desired signal to noise ratio (SNR) or Eb/No. Both of these relationships describe the ratio of the desired signal strength to the strength of noise, both normally measured in dB. The Eb/No equation is given below, where E_b is energy per bit in dB, N_o is noise, P is transmitted power, L_l is transmitter to antenna line loss, G_t the gain of the transmit antenna, L_s is space loss, L_a is the loss due to the transmission path, G_r , k is Boltzmann's constant, R is the data rate, T_s is the receiver system noise temperature.

$$\frac{E_b}{N_0} = \frac{PL_l G_T L_s L_a G_r}{kRT_s} \tag{0.4}$$

For the purpose of this research project all antennae are assumed to be parabolic. It is a very well-known fact that spot beams significantly increase the RF efficiency of a system through frequency re-use as well as better gains. However, in an attempt to reduce the problem set, the research team decided to assume that all the antennae would be parabolic shaped. The gain of a parabolic antenna is a function of the antenna efficiency, the frequency being used and the diameter of the parabolic dish. The equation used to calculate the gain of the antenna is captured in the equation 0.5 below Equation 0.6 is 0.5 expressed in dB. Where G gain, η is antenna efficiency, D is the diameter of the dish, λ is wavelength, f is frequency in GHz, and c is the speed of light.

$$G = \eta\left(\frac{\pi D}{\lambda}\right)^2 = \eta\left(\pi D \frac{f}{c}\right)^2 \tag{0.5}$$

$$G = 20.4 + 20\log(f) + 20\log(D) + 10\log(\eta) \tag{0.6}$$

Because parabolic antennae focus the energy being radiated into a more coherent beam they offer significantly better performance over an omni-directional antennae that radiates in an isotropic pattern. Therefore, by using a parabolic antennae to shape the beam the amount of transmitted power can be significantly reduced if an omni-directional antenna were used. This is not to say that omni-directional antennae are not used in space communications, they have their advantages as well. For example an omni-directional antenna doesn't need to be pointed at the target it's trying to communicate with, however the it requires either increased transmit power, a receive antenna with very high gain or some combination thereof.

The signal that reaches the receiver is significantly weaker than that which is transmitted due to the various losses the signal must endure. By far the most significant loss the signal encounters is the free space loss. Free space loss is calculated in equation 0.7 below, where r is the distance between the transmit and receive antennae. As is evident from the equations, the transmitted power falls off at the rate of $\frac{1}{r^2}$.

$$L_s = \left(\frac{4\pi r}{\lambda}\right)^2 \tag{0.7}$$

The approach the research team took to designing the communications link was to ensure the link could be closed from the ground terminal to the satellite when the distance between them was greatest. This meant that elevation angle was the smallest, referred to as the minimum elevation angle. It is important to note that the power transmitted from both the ground terminal and the satellite is assumed constant and that by designing the link to close at the minimum elevation angle, and therefore the greatest distance, it is assumed that the terminals will have the requisite power to close the link when the satellite is at its shortest distance from the ground terminal. This occurs when the satellite is directly overhead the ground terminal, when the sub-satellite point is the ground terminal or when the elevation angle is 90 degrees. However,

it should also be noted that when the satellite is furthest away from the ground terminal the relative velocity between the two objects is at its smallest resulting in the least amount of Doppler shift experienced by the frequency transmitted. Likewise, when the satellite is closest to the ground terminal the relative velocity is at it's greatest and therefore the Doppler shift experienced by the frequency transmitted is at it's greatest. Therefore, it is assumed that the notional receiver can compensate for the Doppler shift and is able to stay tuned to the correct frequency. It is also assumed that the ground terminal is capable of tracking the satellite, this is normally done by uploading the correct ephemeris files to the ground terminal.

1. Low Earth Orbit

As stated above the first step in designing the link is to determine the distances involved. Figure 13 is used to help illustrate the relationships between the satellite and the ground user.

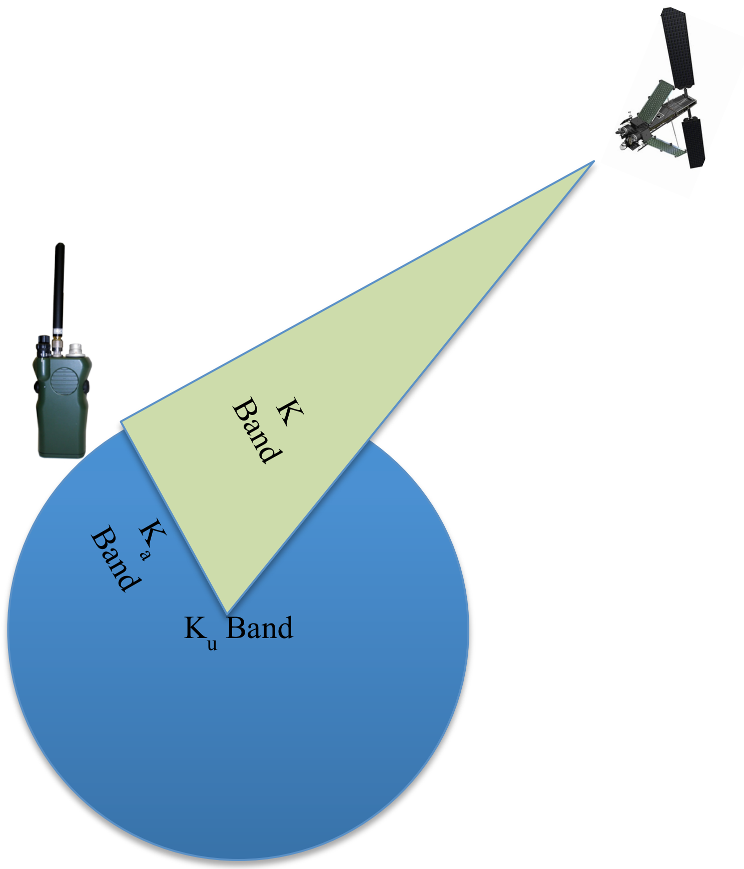

Figure 13. Iridium Satellite Geometry

This represents the moment in time when the satellite is lowest in the sky and can still close the link with the user on the ground. This means the user is on the limb of the earth and transmission path forms a 90 degree angle with the surface of the earth. The first step in the process is to solve for the angular radius of the earth, ρ. This is the longest distance the signal will need to traverse. We can now use the law of sines to solve for the remaining angles. Note that in this particular instance, $\lambda = \lambda_o = 26.84$ and $\rho = \eta = 63.17$ degrees.

(1) Identifying the Hand held and ground terminals . The approach the research team took with regards to designing the communications links was to use existing user terminals. The purpose of going this route was twofold. First, it was the most cost effective for our theoretical constellation. By using existing, primarily software definable radios, there are no costs incurred to

design an entirely new handset to communicate with the satellite network. For the mobile users there are several options that are available, including the AN/PRC 155 networking Manpack Radios produced by General Dynamics, and the ITT NexGen RO Tactical radio.

The AN/PRC 155 is a two channel man portable, software definable radio that supports a variety of waveforms including but not limited to, Soldier Radio Waveform (SRW), Mobile User Objective System (MUOS), and SINCGARS. The AN/PRC 155 is also DAMA compatible making it extremely useful with UHF Follow on (UFO) and the aging FleetSatCom communications satellite constellations.

The ITT NexGen RO Tactical radio is compatible with the Irridium LEO Constellation and is therefore assumed it can compensate for the Doppler shifts associated with communicating with satellites in LEO and traveling at high relative velocities. This radio offers a myriad of features including integrated GPS, Push to talk (PTT), and AES 256 encryption all in a small lightweight hand held form factor. This radio is enabled by the Defense Information Systems Agency's (DISA) Enhanced Mobile Satellite System (EMSS).

While neither radio offers true plug and play capabilities they are without question getting closer and closer to meeting the Military's need for a ubiquitous global communications network.

The ground terminals will consistent with that of the Army's LMST or SWAN which provides a 2.5 meter parabolic dish with a gain calculated at 51.78 dB. The LMST was chosen because is consistent with the size of system used by larger Forward Operating Bases (FOBs) and will be capable of providing the bandwidth required by the significantly larger number of users on a FOB.

2. Defining the Communications Requirements

The first step in designing a communications link is to have a very good idea of the type of information the links will transport and how fast and reliable it needs to get to the intended recipient. For example if your communications link is only required to transmit voice you may not be required to implement any type of compression scheme. Chances are, however, that if your satellite will be communicating with more than one terminal, will only have a relatively narrow slice of bandwidth allocated to it and may be limited in other ways then some sort of compression or encoding scheme will need to be implemented. Additionally if you are designing a link to be used by the military, you will want a signal that has a low probability of detection and a low probability of interception. Additionally you can employ some forward error correction techniques that will reduce the amount of power your terminal is required to transmit but at the cost of a reduced data rate available to the user. These are trades that must be addressed and the requirements known very well up front. Not having these requirements known up front can cause significant problems as the project moves forward. Furthermore, not identifying the trade space will cause significant delays in the project as well.

For the purpose of the research project the research team decided on the following baseline requirements. It should be noted that this is a starting point and that some of the details mentioned above are omitted in an attempt to scope the project to a more manageable scale. These baseline requirements are captured in Table 3.

Communications Links				
Base Line Requirements				
	% Coverage	Data Rate	Freq Band	LPD/LPI
Mobile User	100	500kbps	L	Yes
	100	8mbps	Ka	Yes
Ground Terminal	100	125Mbps	Ka	Yes

Table 3. Communications Baseline Requirements

3. Baseline Link Design Example

The research team initiated this portion of the project by determining the longest distance between the user and the satellite for both the LEO and GEO constellations. Additionally the distances between satellites for both the LEO and GEO constellations were calculated in order to identify the required power being transmitted as well as the required receive gain on the satellite.

Because there is a relationship between the elevation angle and the path length, the path length will increase as the elevation angle decreases. The authors of *An Overview of Iridium Low Earth Orbit (LEO) Satellite System* tell us that the minimum elevation angle for the Iridium constellation is 8.2 degrees (Pratt, et al. 1999). The first step in the process to calculate the propagation length was to calculate the angular radius of the earth which is a function of the altitude of the satellite and will not change as the elevation and nadir angles change. Equation 0.8 describes how the angular radius of the earth was calculated.

$$\rho = \cos^{-1}\left(\frac{R_e}{R_e + h}\right) \tag{0.8}$$

The next step was to calculate the nadir angle. Equation 0.9 describes how the nadir angle was calculated.

$$\sin^{-1}\eta = \cos\varepsilon \sin\rho \tag{0.9}$$

Next, the earth central angle was calculated using equations 0.10

$$\lambda = 90 - (\varepsilon + \eta) \tag{0.10}$$

Now the propagation length can be calculated using equation 0.11

$$D = R_e \sin\lambda \sin\eta \tag{0.11}$$

Now that the distance is known it is possible to determine the characteristics of the receive antenna required to close the link. The first step was to calculate the free space loss. Because this is a function of the distance

the signal must travel it will increase as the distance increases. The space loss was calculated using equation 0.12

$$L_s = \left(\frac{\lambda}{4\pi D}\right)^2 = \left(\frac{c}{4\pi Df}\right)^2 \tag{0.12}$$

Next the atmospheric attenuation must be calculated. In order to do this the atmospheric attenuation at zenith must be determined using Figure 16–18 (from Wertz, Everet and Puschell 2011). The atmospheric attenuation was calculated using equation 0.13. Note this includes a 0.3 dB loss due to polarization mismatch.

$$L_a = \frac{L_{zenith}}{\sin\varepsilon} - 0.3 \tag{0.13}$$

The next variable to be determined is the attenuation due to rain. This was obtained from Figure 13–11 of (Wertz and Larson 2007).

Next the Effective Isotropic Radiated Power (EIRP) was calculated using the equation 0.14.

$$EIRP = L_l + P_t + G \tag{0.14}$$

Next the gain required by the receiver was calculated by using equation 0.15. Where Ts is the noise temperature of the system in kelvin and R is the data rate in bps.

$$\frac{E_b}{N_0} = EIRP + L_s + L_a + L_r + L_p + G_r + 228.6 - 10\log T_s - 10\log R \tag{0.15}$$

Solving for Gr:

$$G_r = -EIRP - L_s - L_a - L_r - L_p - 228.6 + 10\log T_s + 10\log R - \frac{E_b}{N_0} \tag{0.16}$$

The equation for calculating the gain of a parabolic antenna is:

$$G_r = \frac{\pi^2 D_r \eta}{\lambda^2} \tag{0.17}$$

Solving for D_r yields:

$$D_r = \sqrt{\frac{G_r \lambda^2}{\pi^2 \eta}} \tag{0.18}$$

This process was repeated at an iterative process starting at Iridium's minimum elevation angle of 8.2 degrees and increased in increments 0.2 degrees to 90 degrees. The propagation path length and required receive gains are displayed in Figures 16 and 17 respectively.

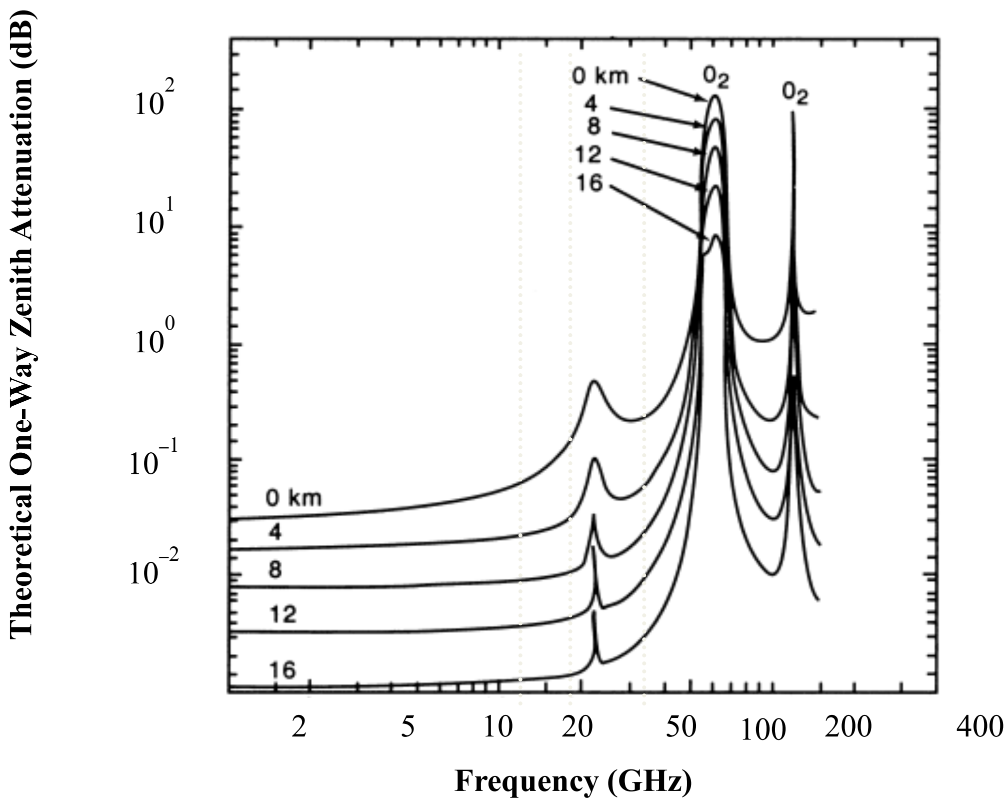

Figure 14. Atmospheric Attenuation at Zenith (from Wertz and Larson 2007)

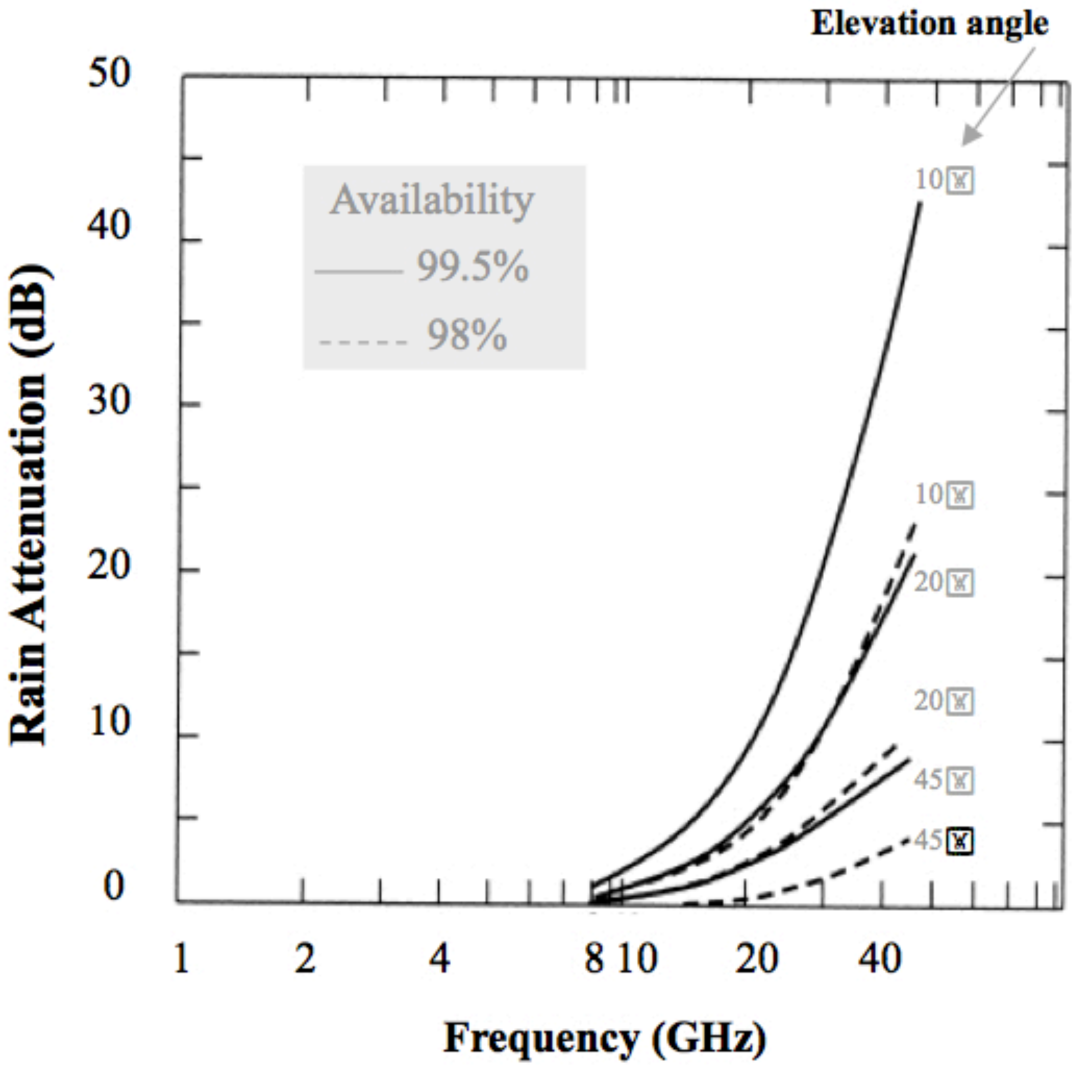

Figure 15. Rain Attenuation (from Wertz and Larson 2007)

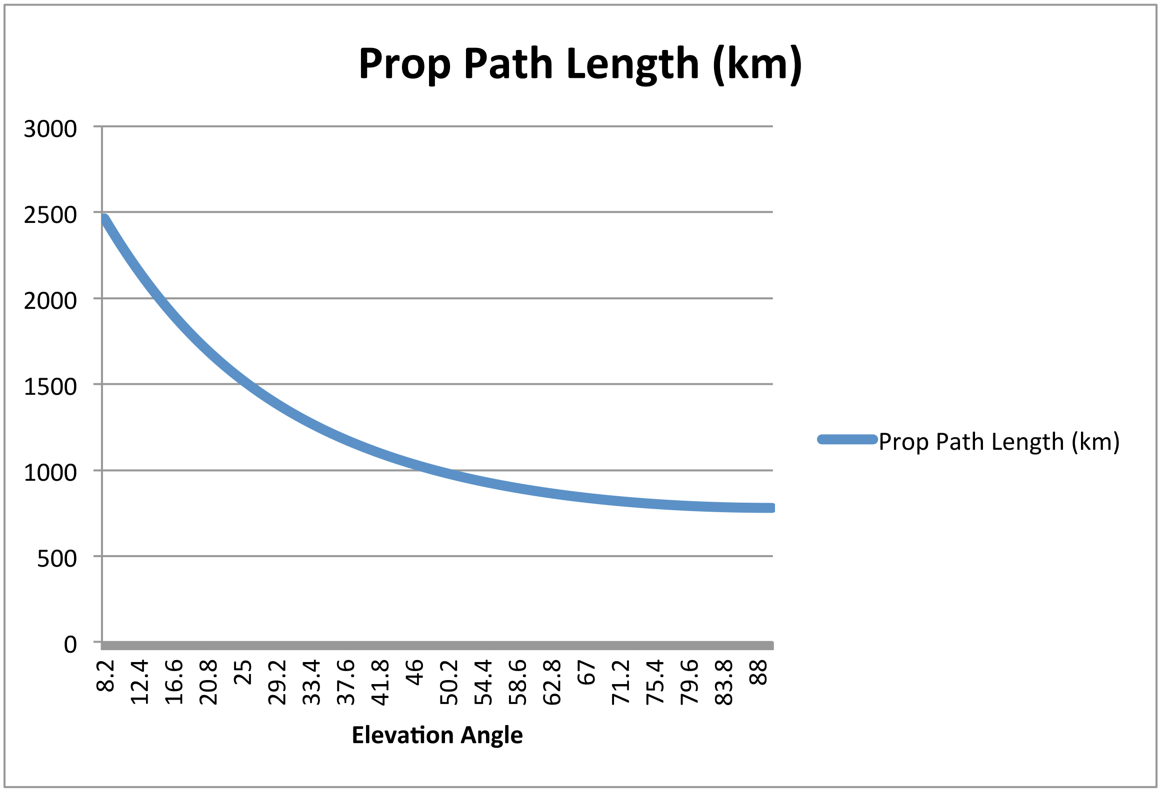

Figure 16. Propagation Length As A Function of Elevation Angle

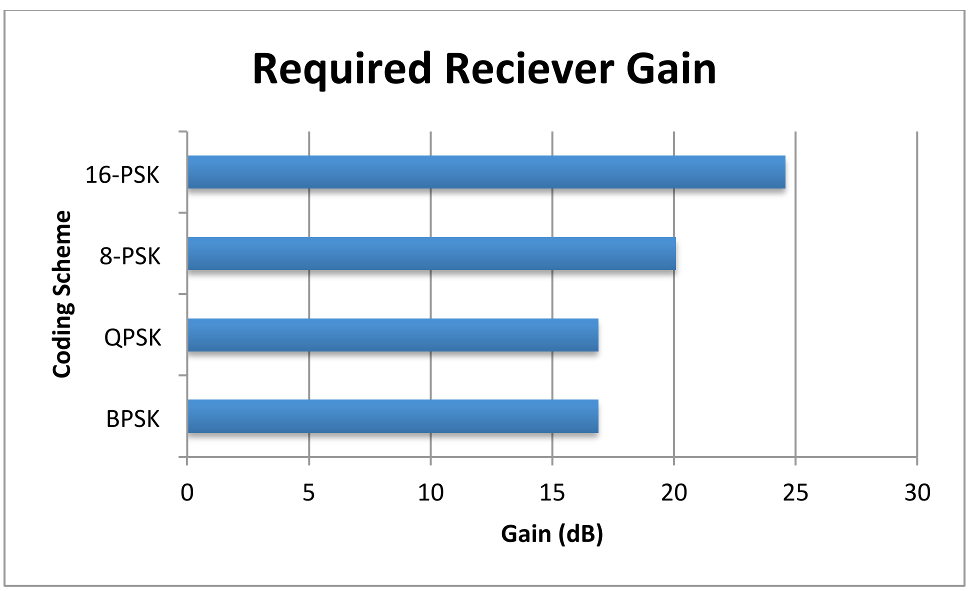

Figure 17. Required Receiver Gain

Because the Iridium Next constellation advertises a future Ka band capability the same calculations were performed. However, there were some changes that needed to be made. For Ka the frequency used was 20 GHz, the transmit antenna was 6 inches in diameter, output power was 20 watts, the atmospheric attenuation was -1.3 dB the rain attenuation was -3 dB. The required receiver gain and required receive antenna diameter are captured in Figures 18 and 19.

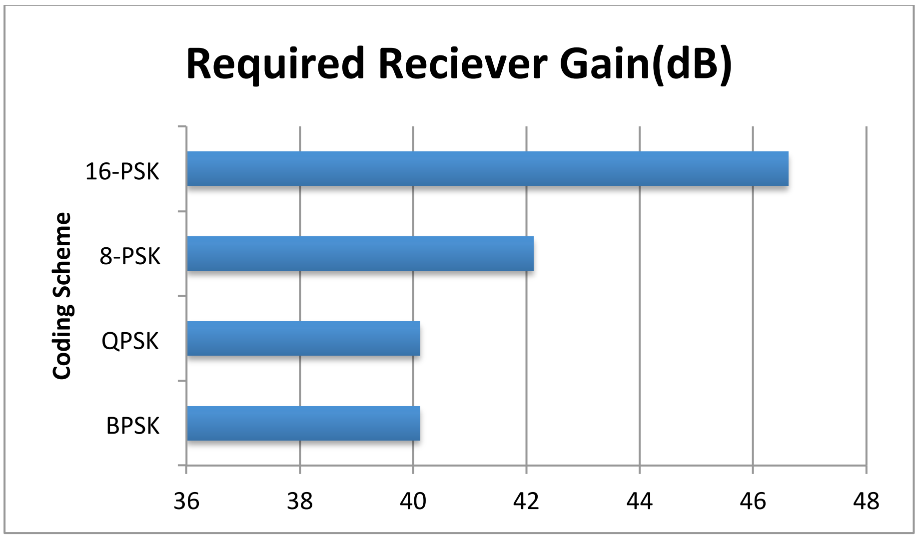

Figure 18. Required Ka Band Receiver Gain

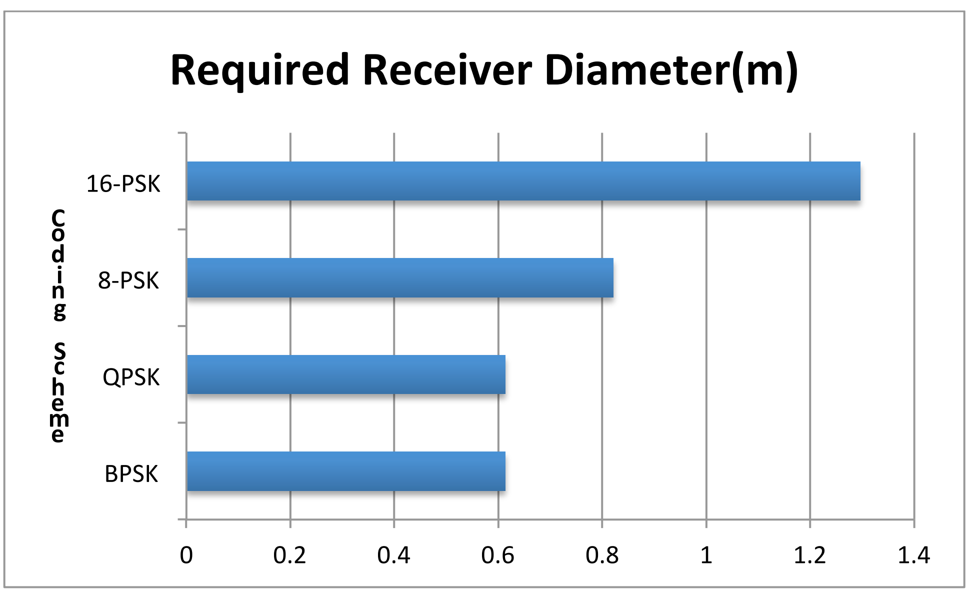

Figure 19. Required Receive Antenna Diameter

The receive antenna diameters displayed in Figure 17 are assumed to be parabolic dishes. It is important to note that the required receiver gains displayed in Figure 16 are un-coded. Meaning, there is no forward error correction (FEC) scheme implemented. FEC encoding schemes have the ability to reduce the probability of errors and even correct errors that occur during transmission, to a degree. Some error correction schemes can correct more errors than others. This does not come without a cost. In a fixed bandwidth environment, like most military communication systems, an error correction scheme means a reduction in the data rate the user or terminal has access to. This is because there are extra bits that are transmitted and these bits are not part of the data packets providing the information the terminal or user requested. The benefit of error correction schemes is that they reduce the amount $\frac{E_b}{N_0}$ required to close the link. This in turn translates into a reduction in the receiver gain and a reduction in the required diameter of the receive antenna. While these reductions may appear to translate into an overall cost savings for the entire system, that may not be the case. In some instances, money saved by reduction the size of the antenna may need to be redirected to increase to computing capacity of the satellite to ensure it is capable of performing the calculations quickly and accurately enough to provide a reliable communications system.

4. Doppler Shift

Doppler shift is described as the change in the observed frequency due to speed of the transmitter relative to the receiver, the change in the observed frequency due to the receiver relative to the transmitter. You may have noticed the change in the pitch of a car as it passes you while you are standing on the side of the road. As it approaches you the sound waves it produces are getting increasingly bunched up, resulting in the frequency increasing. As the car passes you, the sound waves are increasingly getting spread out, resulting in a lower pitch.

This change in frequency could prove to be disastrous for your system if it is not capable of compensating for it. A communications system must be capable of measuring the range from transmitter to receiver. Satellites accomplish this by transmitting pseudo random codes or tones. These are received by the satellite (and ground terminal) and are then transmitted back. Once received, the system can then calculate the range and the range-rate to determine the relative velocity and ultimately the shift in frequency.

Satellite Tool Kit was used to model all the satellites, ground stations and communications systems studied by the research team. The assumption was made that the Doppler shift experienced by one of Iridium's satellite would be the same by all of the Iridium satellites. A scenario was created with just one Iridium satellite. A detailed link budget report was then generated to capture the information about the communications link, including, but not limited to Doppler shift. Because this particular scenario was engineered to illustrate only the Doppler shift there were no constraints, such as bit error rate (BER), $\frac{E_b}{N_0}$, or data rate placed on the link. The scenario ran for a 24 hour period. Figure 20 illustrates how the scenario was set up. Figure 21 illustrates the accesses the ground station has to the satellite. Figure 22 illustrates the change in frequency. Figure 23 illustrates the effects of Doppler shift has on the received frequency. Figure 24 illustrates the rate of change of the received frequency. The rate of change is extremely important because it will become a design constraint for a future system. It is important to note that the frequency used to illustrate the effect of the Doppler shift was a Ka band frequency, 20 Ghz.

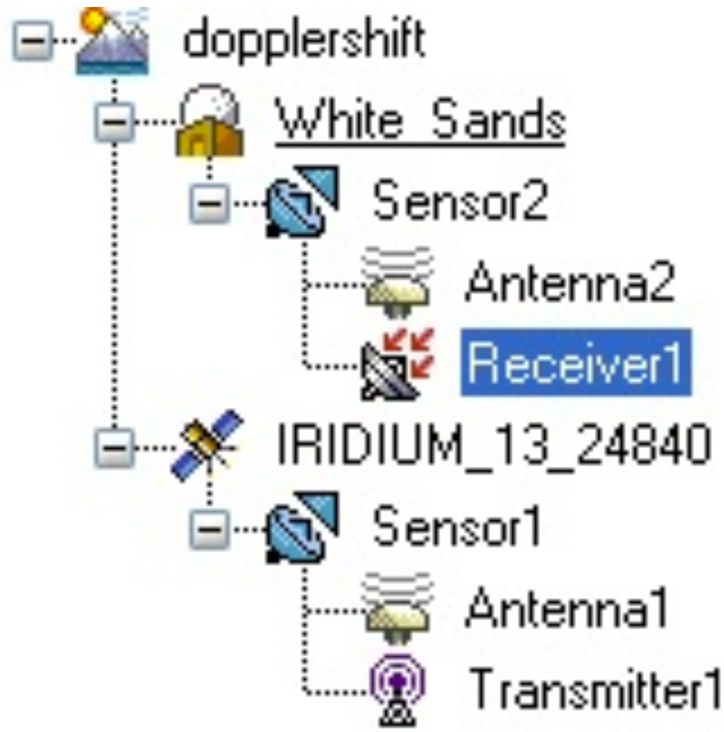

Figure 20. STK Scenario Setup

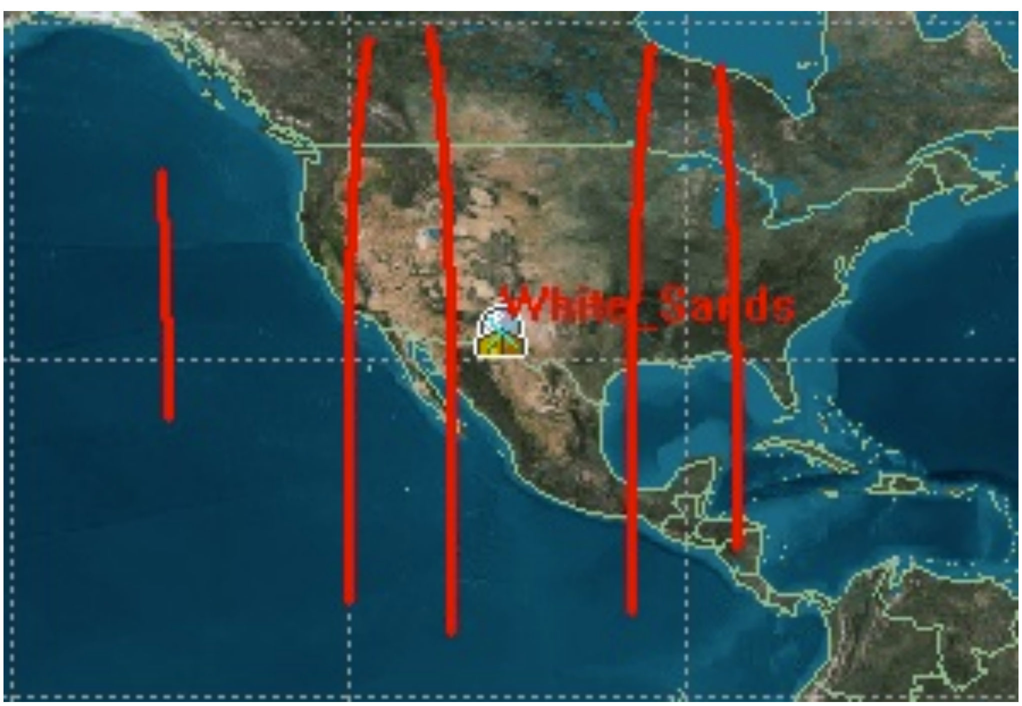

Figure 21. Satellite Accesses

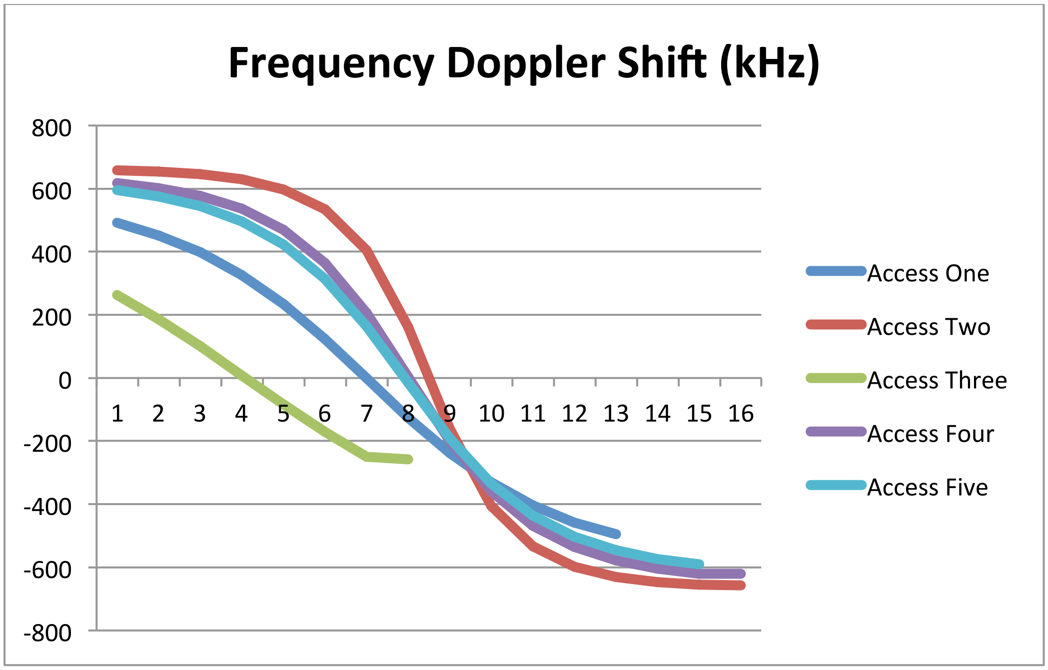

Figure 22. Doppler Shift

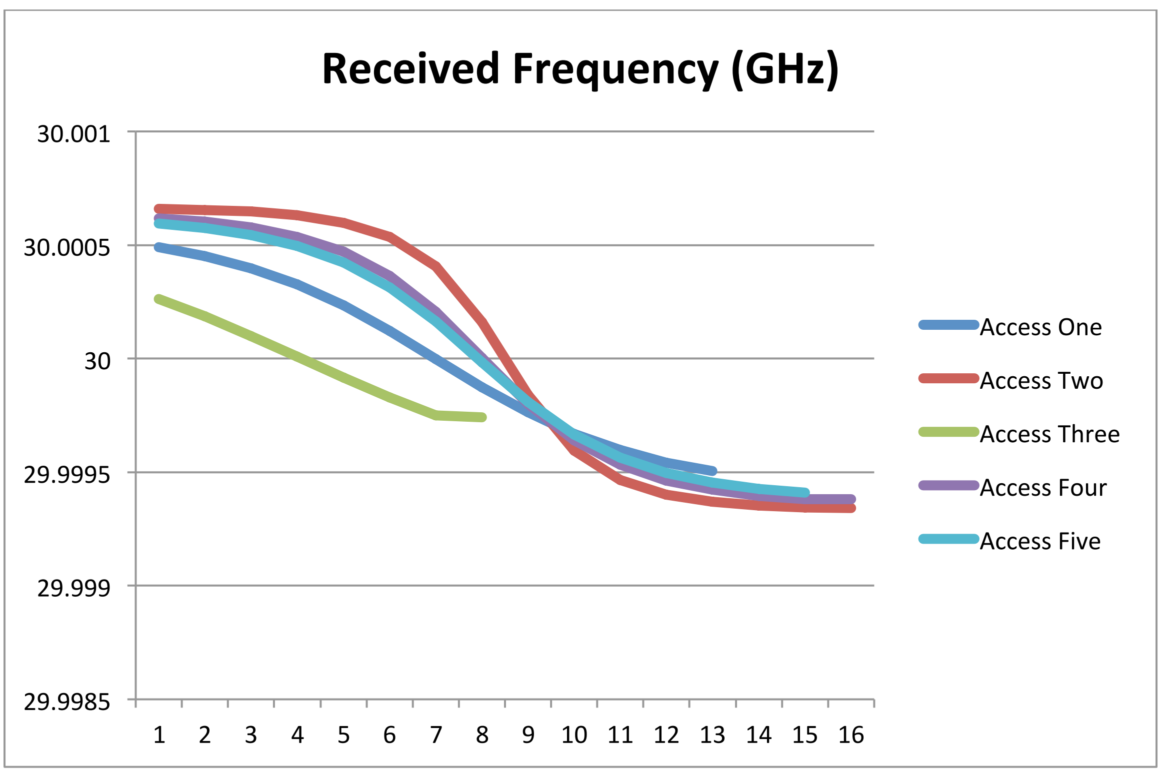

Figure 23. Received Frequency

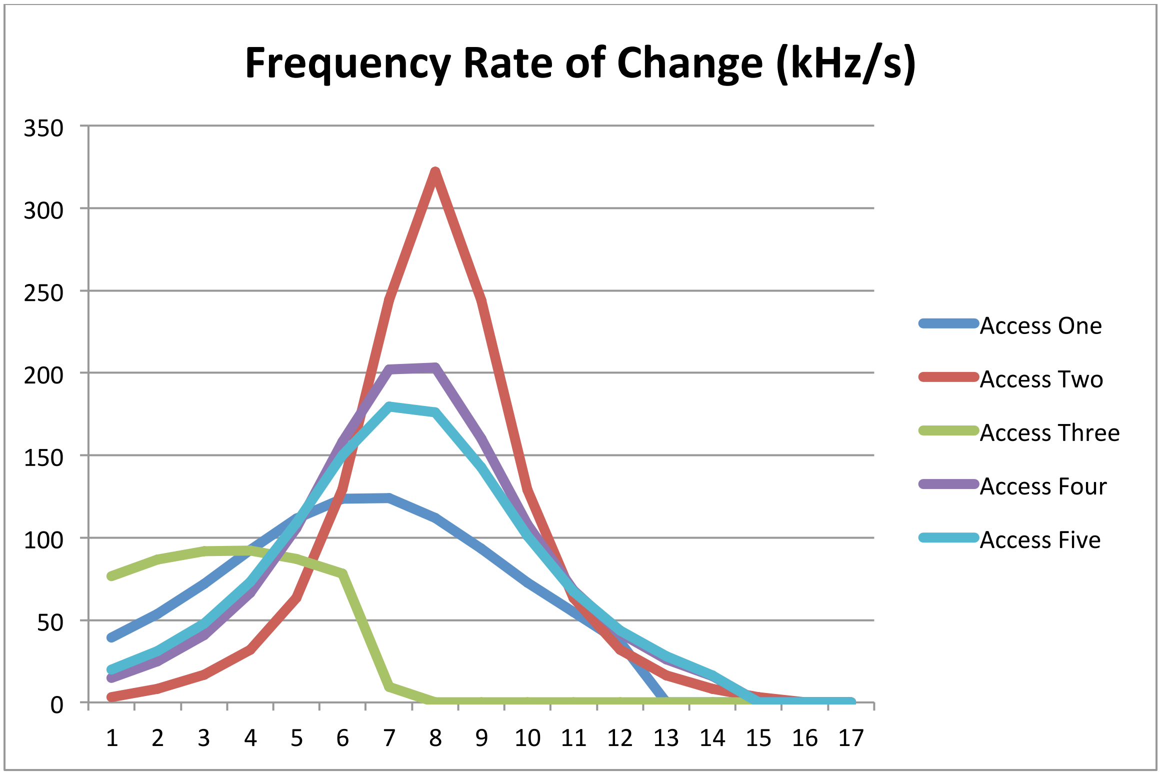

Figure 24. Frequency Rate of Change (kHz)

B. GEOSYNCHRONOUS ORBIT.

Because the orbital period of a geosynchronous orbit is equal to amount of time it takes for the earth to revolve on it's axis (one sidereal day), a geo satellite is always looking at the same spot on the earth. Depending on the inclination of the satellite it may drift above or below the equator as the satellite progresses through its orbit. Relative to the earth, a geo satellite is stationary and thereby greatly reducing the complexity of the communications link design because the problem set is reduced. Because the satellite is stationary there is virtually no Doppler shift to compensate for and essentially reduces the problem to power and gain levels to ensure the satellite can still close the link to a ground station or user on the limb of the earth thereby making full use of the satellites foot print.

THIS PAGE INTENTIONALLY LEFT BLANK

V. SATELITE NETWORKING TEST BED INTEGRATION

A. MODELING AND SIMULATION

The inherent complexity of the various phenomenology at work within both satellite communications architectures and within information networks drives the need for robust modeling and simulation. Modeling and simulation however, has several dimensions of performance that must be considered before being used to analyze a real world complex system. The first dimension of performance is fidelity, which is a measurement of how accurately the model reflects reality. Achieving a high degree of fidelity within a model becomes increasingly difficult as the complexity of the system to be modeled increases. The second dimension of performance concerns how the model deals with the progression of time or events. Two main types are discrete or continuous models. Discrete models deal with discrete events or commands as the inputs to the modeling system. A special subset of a discrete model is a discrete event simulation which moves the model through time at some pre-determined time step which considers the model as a static entity for the duration of that time step. A continuous simulation performs numerical solutions to differential equations and uses the solutions as inputs for further iterations of the model. The third measure of performance of a model is resolution, which deals with the level of complexity of the modeled environment and the granularity of its associated data or with the type of environment being modeled (Rainey 2004).

In order to model and simulate the behavior of a space based information network, a high fidelity, discrete event, multi-resolution model must be developed. The requirement for fidelity is self-explanatory. A discrete event simulation allows the model to progress gracefully through time, allowing for human cognition to participate in the analysis since the model behaves as if it were real life and can be observed as such. Multi-resolution modeling is building a single model or a family of models that inform at various levels or complexity or

within several interrelated but distinctly separate environments (Davis and Bigelow 1998). In this case, the physical laws which determine the behavior of a constellation of satellites in orbit around the Earth communicating via electromagnetic waves are distinct from the physical laws which govern the communication of nodes within the information network overlaid on that satellite constellation. The two systems are interdependent and yet distinct. This drives a need for a comprehensive modeling methodology that allows the interplay of these two systems in a discrete event simulation following a single time step and common modeling environment.

An additional motivation for multi-resolution modeling is that complex adaptive systems often exhibit emergent behaviors, or regular coherent behavior at the macroscopic level that are not readily apparent or understandable in terms of the microscopic laws that govern the system. These emergent behaviors make complex adaptive systems both difficult to experiment with and oftentimes counterintuitive at the macro level. Multi-resolution modeling allows for lower resolution models to be combined in a way that preserves the emergent behavior exhibited at the macro level without sacrificing the fidelity of the model as a whole. High resolution models tend to require such complexity that their scope must be constrained, making them unwieldy for experimenting with emergent phenomena in complex adaptive systems. In this way, it is preferable to use multiple specialized modeling tools aggregated into one multi-resolution model.

In the case of a satellite based information network, there are two primary focuses of modeling. The first is the physical environment which includes the satellite and ground station “nodes” that make up the constellation and their motion as well as the electromagnetic medium used for communication. The relative positions of each satellite must be modeled in order to understand the changing accesses over time. The relative velocities determine the amount of Doppler shift inherent in the communications link. The second environment is the network topology, which relies upon the finite state of the physical model at any

given time. Into this discrete physical network topology, the rest of the environment deals with various settings at each of the levels of the OSI model.

This modeling environment can be visualized as several individual models aggregated together in a hierarchy. The physical location and movement vectors of the nodes comprise the first level in the hierarchy. Onto this, the communication links must be modeled using the position vectors as inputs. This communication model displayed in Figure 25 provides the foundation for the network model by providing the communications link metrics as inputs to the network model.

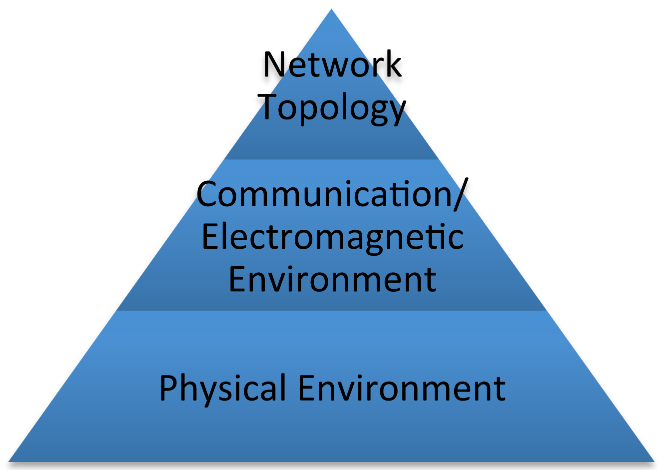

Figure 25. Model Aggregation Hierarchy

In order to create a modeling environment with the desired characteristics within the timeframe available, a commercial off the shelf solution is preferable. Given the importance of the physical environment modeling, that is the type of modeling software that we focused on. So the requirements for the nascent modeling environment software are a discrete event physics based modeling and simulation program with a robust capability for extensions into other forms of modeling such as communications and networking. We desired to have a single modeling environment with a common graphical user interface (GUI). As a result, we chose Analytical Graphics' (AGI) Satellite Tool Kit (STK) for modeling the

physical and communications environments and Qualnet to model the network behavior. Qualnet was chosen specifically due to its ability to interface with STK and run within a common GUI.

B. PHYSICAL ENVIRONMENT MODELING

The overall design philosophy in modeling the physical environment was to mitigate any network errors that might occur due to the physical environment and as such, the attributes of the physical layer of each satellite network are designed to limit the possibility of lost data due to poor link quality. The purpose of the modeling effort is not to define the best fit protocol for a specific satellite constellation, but rather to discover the performance of different protocols across several generic constellations. So by removing, to such extent as is possible, the potential for lost data on the physical layer, it can be assumed that any discrepancies in performance between protocols is due to the behavior of the protocol and not the performance of the constellation.

1. Satellite Tool Kit

Satellite Tool Kit is a physics-based software package from Analytical Graphics, Inc. that enables scientists, engineers, and researchers to perform complex analysis of spacecraft mission design and operations, space exploration, communications analysis, C4ISR (battlespace management), electronic warfare, geospatial intelligence, unmanned systems (UAVs), and missile defense. The core of STK is a geometry engine that determines the time-dynamic position and attitude of objects, determining dynamic spatial relationships among all of the objects under consideration including the quality of those relationships or accesses given a number of complex, simultaneous constraining conditions. STK has been developed over 20 years as a commercial off the shelf software tool. Originally created to solve problems involving Earth-orbiting satellites, it is now used in both the aerospace and defense communities (Analytical Graphics Inc. 2012).

STK began in the aerospace community for orbit analysis and access calculations, but as software matured, more modules were added including the ability to analyze communications systems, interplanetary missions, radar, and orbit collision avoidance. The ability to visualize scenarios in full 3D led to the adoption of STK by military users for real-time visualization of air, land, sea and space assets. STK has also been used by various news organizations to graphically depict complex events to a wider audience.

The interface to STK is a standard graphical user interface (GUI) display with customizable toolbars and dockable maps and 3D viewports. All analysis can be done through mouse and keyboard interaction.

In addition, there is a scripting interface named Connect that enables STK to act within a client/server environment (via TCP/IP) and is language independent. Users on Windows have the option of using STK programmatically via OLE Automation. STK also comes with a “button” tool which uses comma delimited input or excel spreadsheet formatted text commands to perform any of the actions available through the keyboard and mouse interface. The button feature is helpful in streamlining the process for building complex models without divergence due to human errors in data entry.

Each analysis or design space within STK is called a scenario. Within each scenario any number of satellites, aircraft, targets, ships, communications systems or other objects can be created. Each scenario defines the default temporal limits to the child objects, as well as the base unit selection and properties. All of these properties can be overridden for each child object individually, as necessary. Only one scenario may exist at any one time, although data can be exported and reused in subsequent analyses.

For each object within a scenario, various reports and graphics (both static and dynamic) may be created. Relative parameters, between one object and another can also be reported and the effect of real-world restrictions (constraints) enabled so that more accurate reporting is obtained. Through the use of the

constellation and chains objects, multiple child objects may be grouped together and the multipath interactions between them investigated.

STK can be embedded within another application (as an ActiveX component) or controlled from an external application (through TCP/IP or Component Object Model (COM)). Both integration techniques can make use of the connect scripting language to accomplish this task. There is also an object model for more “programmer oriented” integration methodologies. STK can be driven from a script that is run from the STK internal web browser in the free version of the tool. To control STK from an external source, or embed STK in another application requires the STK/Integration module (Analytical Graphics Inc,. 2012).

2. Modeled Constellations

Each of the four constellations modeled were chosen to be representative of a different implementation of a communications architecture. The GEO model is representative of a constellation whose relative complexity is low and which is highly static. The LEO model is representative of a more complex constellation than the GEO and with less static access, but whose complexity is mitigated by the fact that each LEO plane was designed to allow for inter-planar links to operate in a relatively static nature. The Hybrid constellation is a combination of the LEO and the GEO. The Hosted constellation is taken from actual launch history and is the most complex and most random implementation, representative of an ad-hoc network built up over time vice a designed architecture.

a. Geosynchronous

The GEO model is the simplest and most static of the four constellations modeled and is comprised of 4 geosynchronous satellites and four ground stations. The four satellites are generic satellites placed 90 degrees apart and the ground stations were input using STK’s city database and chosen to

make sure that each ground station would be in view of two satellites. This configuration is displayed in Figure 26.

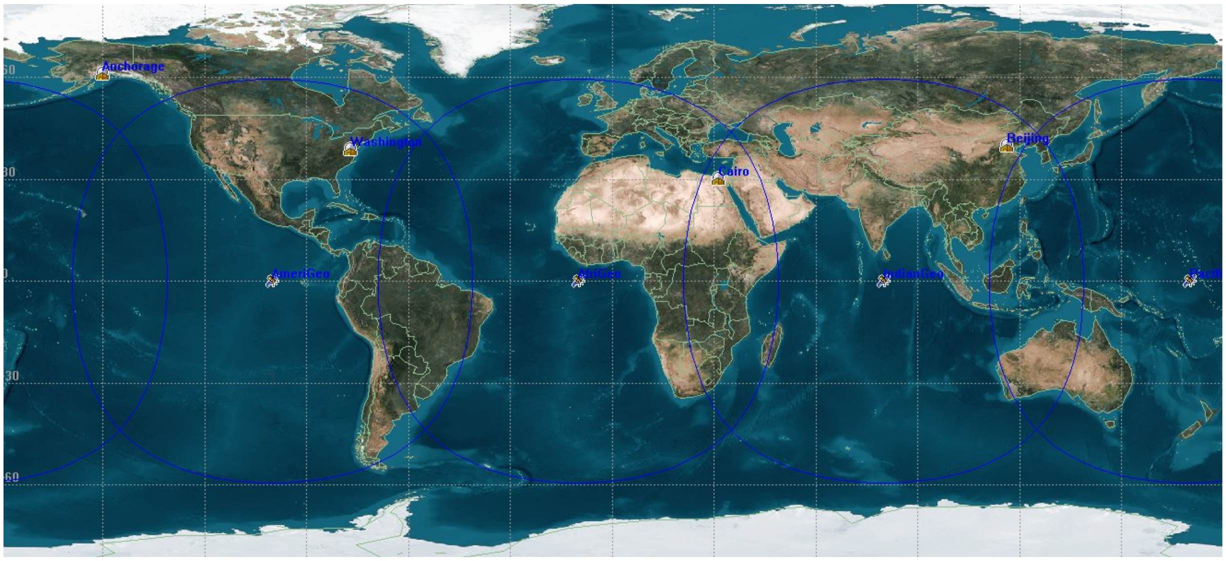

Figure 26. GEO Model Satellite and Ground Station Placement

The communications links, designed to give the greatest assurance of physical layer convergence, are all in the KA band at a 30GHz carrier frequency using antennas which are targeted. The antennas are parabolic dishes. The ground stations have 5 meter dish antennas and the satellites have .7 meter diameter dishes. The antennas in STK are targeted by attaching them to a sensor object and then defining the pointing characteristics of the sensor object. In the GEO model, each ground station has two sensors. Each one is set in targeted mode and assigned to target one of the two satellites in view. The ground station antennas are then attached to the sensors. The satellites have 5 sensors each: an uplink sensor which points directly at nadir, two crosslink sensors which are targeted at the satellites to the east and west, and two downlink sensors which are targeted at the ground stations in view. No communication links were modeled directly between the ground stations. Figure 27 is a wire diagram of the GEO scenario.

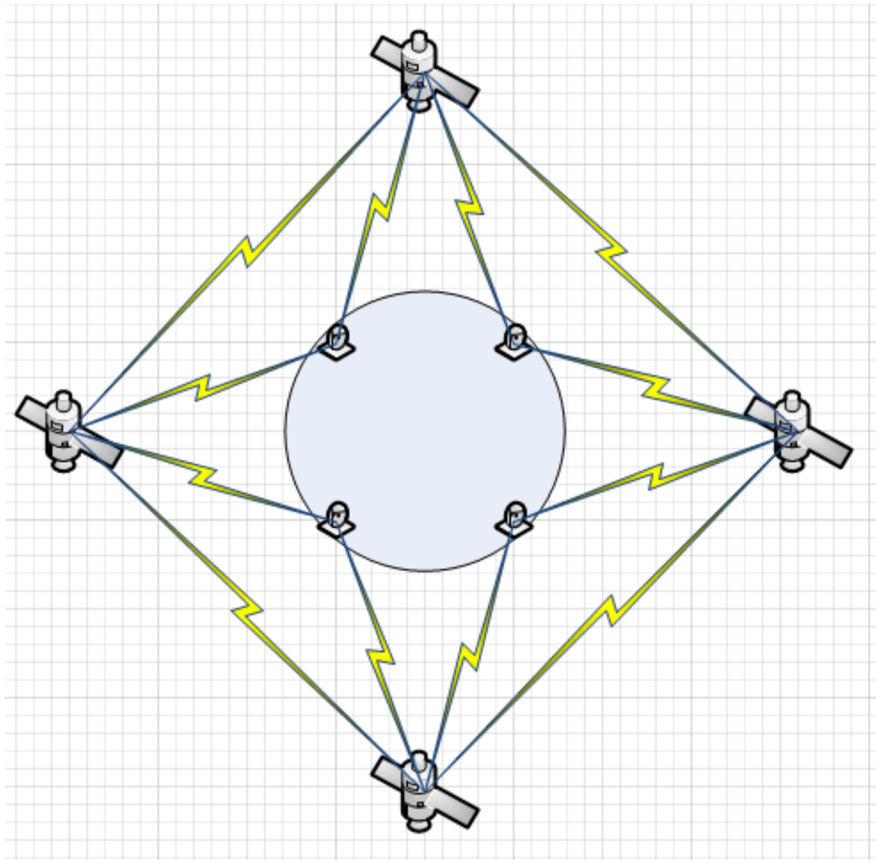

Figure 27. GEO Network Topology

The result is a relatively simple network of four satellites with four links each and four ground stations with two links each. The network topology is temporally static and should give results between the various networking protocols similar to terrestrial networks which are static.

b. Low Earth Orbit

The goal of the LEO model was to simulate a constellation with many nodes and complex orbital motion, but one in which the relative motion between satellites was minimized as much as possible. The best example of a satellite network such as this is the IRIDIUM constellation, which was imported into the model from the STK database.

The Iridium constellation consists of 6 planes of 11 satellites each in an Adams/Rider formation of optimally phased polar orbiting satellites

designed to minimize the total number of satellites required to attain global coverage. Their altitudes are all around 780KM. Since the satellites move in the same relative direction along a common co-rotating interface, the number of connections within the constellation which are static is much higher than those which are dynamic. The dynamic links all exist in one of two places. The first is along the seam where one plane is ascending and the next is descending, and the second is between the constellation and the ground (Adams and Rider 1987). Figure 28 is a 2D representation of the LEO constellation.

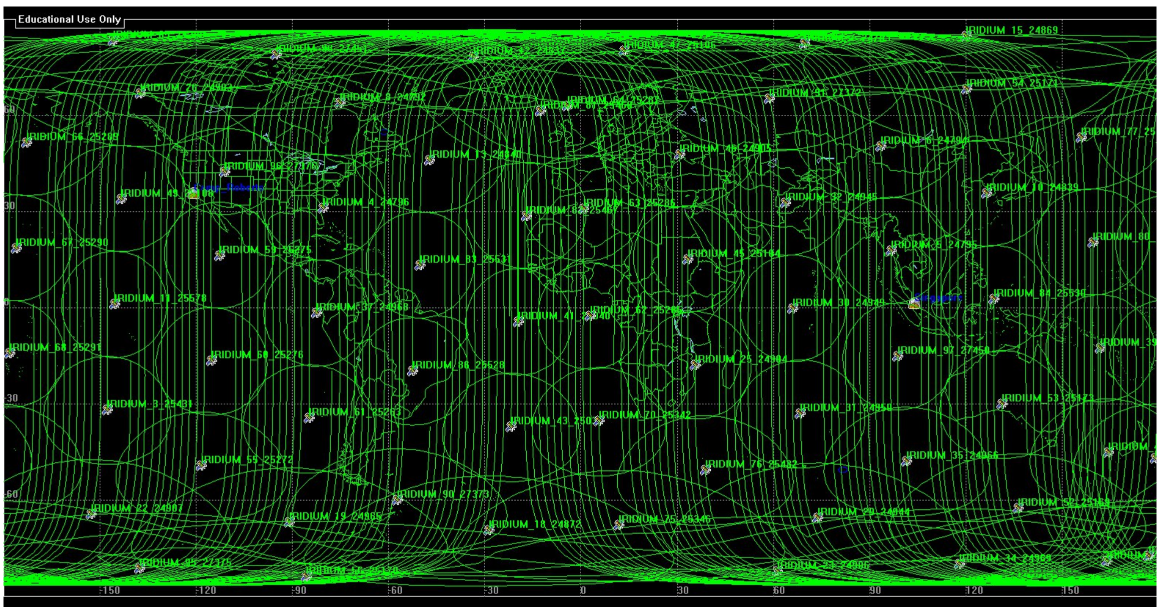

Figure 28. LEO Constellation 2D ground track

The ground stations used, due to their involvement with the NPS Center for Network Innovation and Experimentation (CENETIX), are Camp Roberts in central California, and Singapore. These locations were added to the model using the geodetic reference using latitude and longitude. Each ground station has a single sensor object with is set to target the nearest LEO satellite in view. To this sensor, a parabolic dish antenna with a 30GHz carrier frequency and a 2 degree beam width was added for the ground station uplink.

Each of the satellites has five sensors. One sensor is pointed directly at nadir with a parabolic dish antenna utilizing a 30GHz carrier and 60 degree beam width, providing the downlink. The other four sensors provide crosslinks to the leading satellite ahead in the orbital plane, the trail satellite behind in the orbital plane, the nearest satellite in the plane to the east, and the nearest satellite in the plane to the west. Figure 29 is a 3D representation of the LEO constellation.

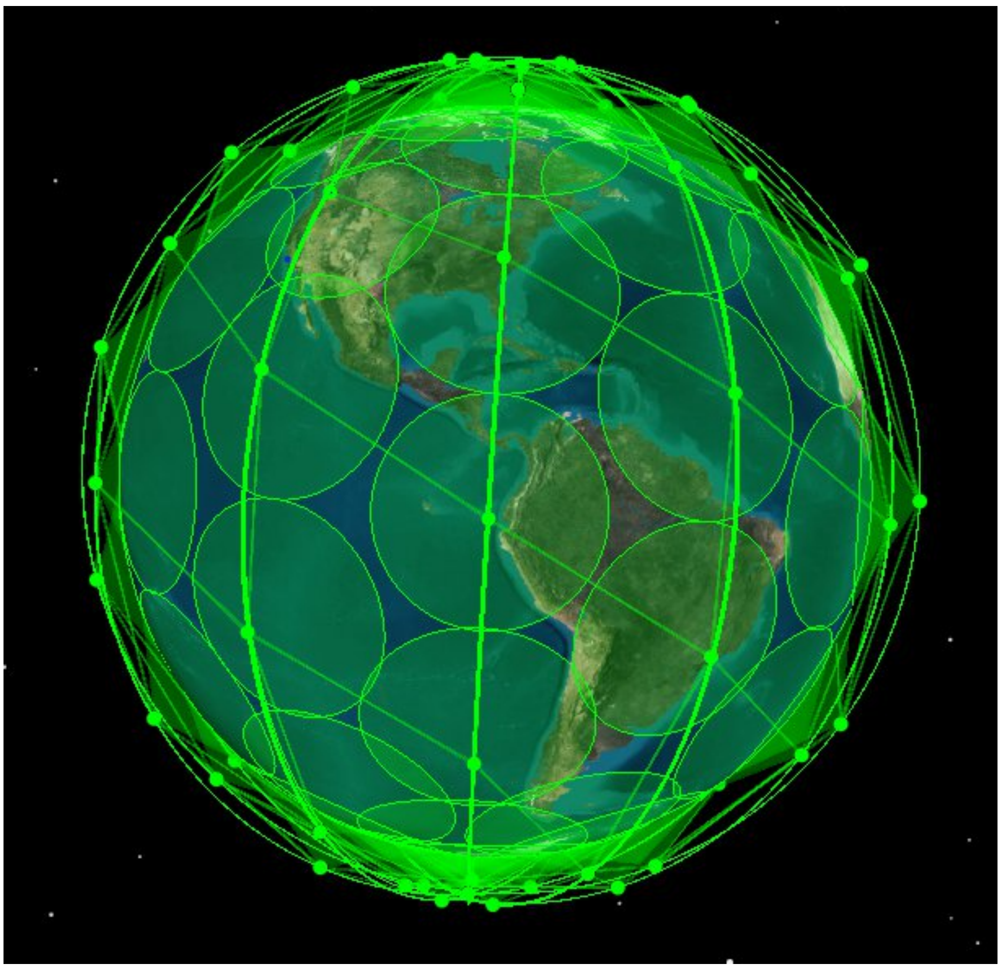

Figure 29. LEO Constellation Crosslinks

These spatial relationships do not change very much throughout the course of the orbit. The exception is on the seam where satellites in one

plane are descending and satellites in the other plane are ascending. Conversely, the node which is connected to any one ground station at a time is constantly changing. It is similar to a cell phone call being passed from tower to tower while the user drives down a highway, except in this case the user is stationary (or relatively so) and the towers are moving at 7KM/sec! As a result, the network topology within the constellation is fairly static, but the interfaces between the constellation and the ground are highly dynamic. Figure 30 is a wire diagram of the LEO network.

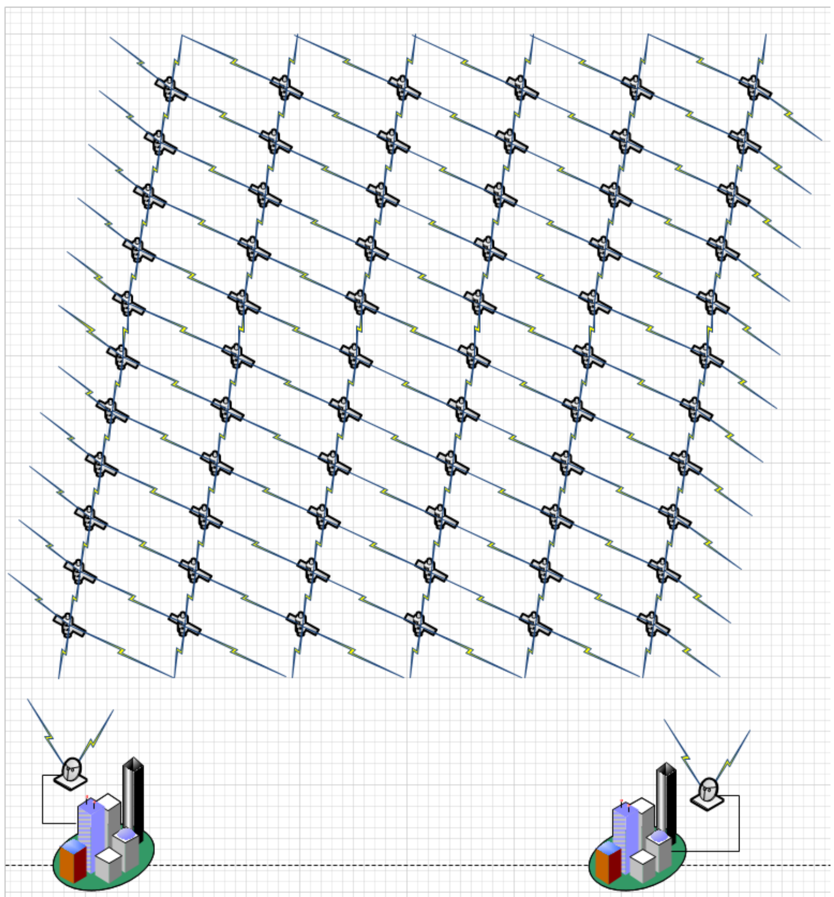

Figure 30. LEO Constellation Network Topology

As a network, the constellation can be visualized as a kind of grid with 6 columns and 11 rows and all of the crosslinks are static except those on the left and right of the grid where the ascending/descending node seem occurs. The ground stations connect to this grid via whichever node happens to be

passing overhead at that time. In order to make routing decisions from one ground station to the other the network must be capable of anticipating which satellite will be in view of the ground station next and be prepared to shift traffic when that happens.

c. Hybrid

The hybrid model is a combination of the LEO and the GEO constellations. The intent is to test the impact of combining highly dynamic and highly static constellations on the various networking protocols. Figure 31 is a 3D representation of the Hybrid network.

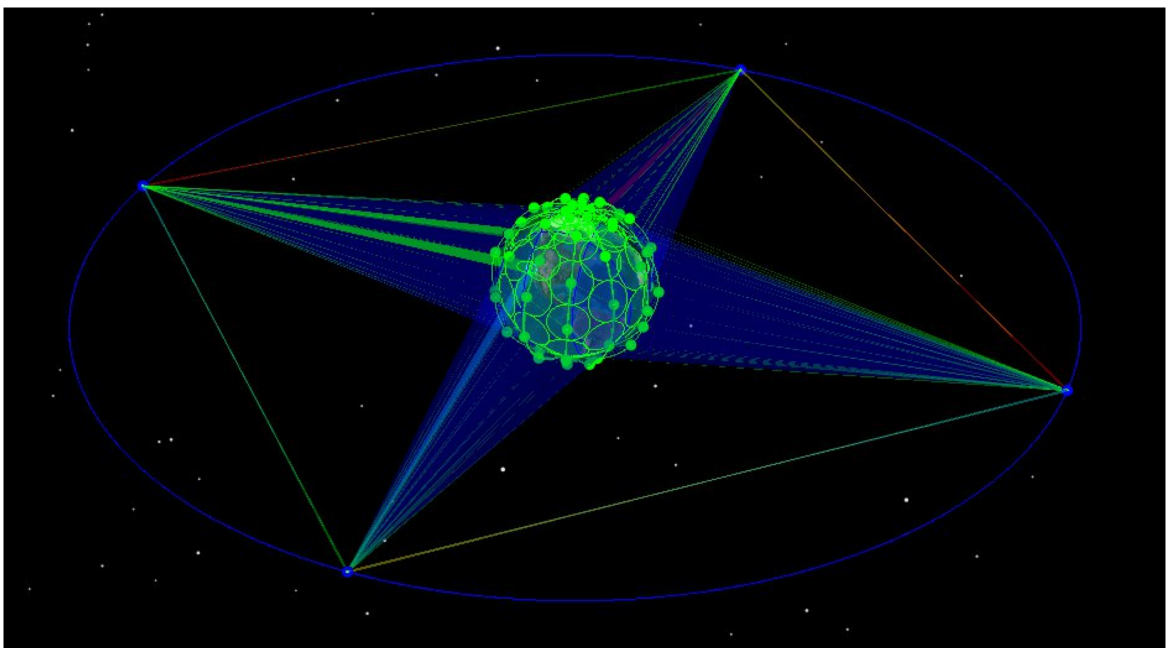

Figure 31. Hybrid Constellation 3D view

The ground stations used were the same as the LEO constellation, Camp Roberts CA, and Singapore. Each ground station was given a sensor pointed at each GEO satellite in view with a 5m parabolic dish antenna. In addition, each ground station was given an antenna targeted to provide the uplink to the closest LEO satellite in view. Each Geo satellite was given the same

crosslinks as in the GEO model, with the addition of a single downlink beam for communicating with the LEO satellites which is pointed directly at nadir. Each LEO satellite has the same communications links as described in the LEO model, with the addition of an uplink targeted to the nearest GEO satellite with the same antenna as the LEO model crosslinks.

In order to visualize the Hybrid constellation, imagine 3 distinct spheres nested within each other. The smallest sphere in the middle is the Earth with the ground stations being the nodes on this sphere. This sphere rotates about the pole. Around this inner sphere is the LEO constellation sphere which is comprised of 6 independently rotating planes which travel in ascending (northward) and descending (southward) directions. The outermost sphere is the GEO constellation which rotates at the same radial velocity as the Earth such that the satellites in GEO appear to be co-rotating with the earth.

As a network diagram, for simplicity, each of these spheres can be represented as an individual network. The interfaces between each of these three spheres have varying levels of complexity and dynamism. For instance, the interfaces between the ground stations and the GEO satellites are static; however the interfaces between the GEO satellites and the LEO satellites are more dynamic, changing twice per LEO period or about every 45 minutes. The interfaces between the ground stations and the LEO satellites are the most dynamic, changing every few minutes as one satellite goes beyond the horizon and another replaces it. Figure 32 is a wire diagram of the Hybrid Network.

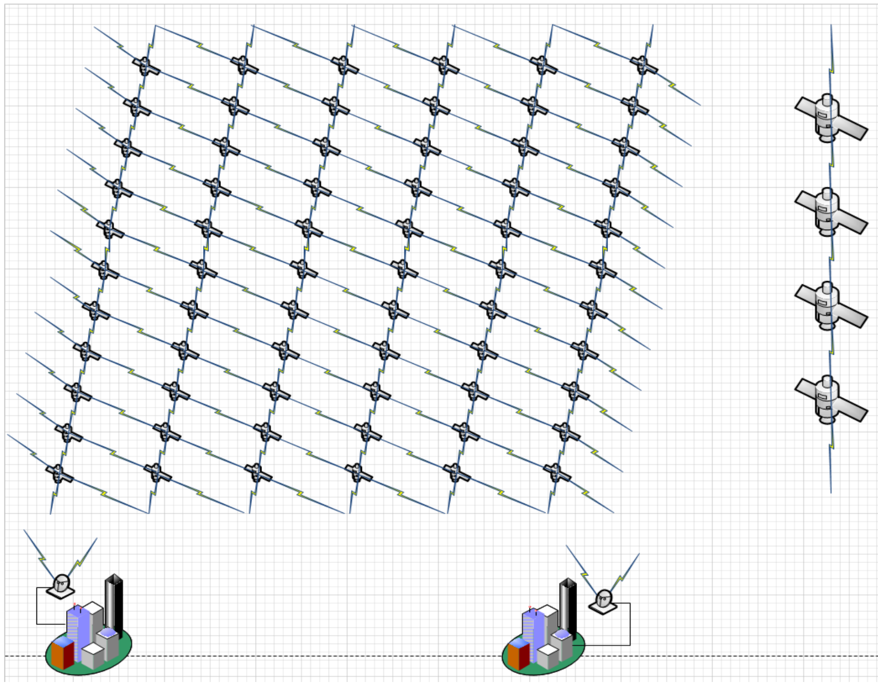

Figure 32. Hybrid Constellation Network Diagram

d. Hosted Payloads

Developing and implementing a model for a routed space-based network will require a major investment in new systems. It is possible that a nation, group of nations or even a private company could launch a satellite constellation capable of creating such a mesh network. By adding space routers on satellites already being launched, this space network could be built gradually. Such a gradual buildup of satellites into a network would result in a network that conforms to no central planning methodology but would instead seem quite random. The hosted constellation allows for testing in a randomized constellation, but in order to be true to life, the satellites chosen must be representative of what a real hosted payload network would look like.

Careful consideration and analysis must be conducted before choosing a satellite to host a routed communications payload. Before any analysis of potential cost savings or network capacity can be accomplished, the criteria for choosing a satellite to host an IP router must be clearly defined. The most important criteria that any potential host spacecraft must meet are the size,

mass and power requirements of the router payload. IRIS has a mass of 90 kg, requires 250 watts of power and 60 Mbps aggregate bandwidth across all channels. If a host spacecraft can't meet these requirements, it cannot be considered for hosting the current IRIS payload. If however IRIS was to become smaller, then the pool for host spacecraft gets bigger (Orbital, 2012).

After determining the general requirements to host the IRIS payload, the type of satellite or rather the mission of the host spacecraft should be considered. While nearly any satellite could potentially host the payload, only communications satellites have the required infrastructure required by the payload already onboard. Communications satellites have the transmitters, receivers, amplifiers, antennas and other ancillary equipment required to communicate with multiple users at the same time therefore not requiring any major changes to the satellite or its mission or the components of the satellite. For example, while an earth observation or imaging satellite may be capable of carrying the IRIS payload into orbit, it probably only communicates with one user at any given time and would therefore require significant changes to the satellite before it would be able to fully support IRIS. Additionally, imaging satellites generally aren't flown in constellations. While more than one imaging satellite may be flown by the same company, DigitalGlobe for example, they are not configured with the cross links required to facilitate communication between the satellites.

The final attributes of potential host spacecraft that need to be analyzed are the orbit and the number of satellites in the constellation. In the case of geostationary satellites, you generally want a constellation with at least two, and preferably three satellites to provide continuous, worldwide coverage.[8] However, for low earth orbit (LEO) satellites, the number of satellites required to achieve continuous worldwide coverage increases significantly.

The analysis began by examining military and commercial communication satellite launches from 2006 to 2011 by the U.S., UK and AUS (Australia did not launch any in this period). These satellites will be the

foundation of the Hosted Payload Constellation. They represent a wide variety of orbits and capabilities ranging from LEO to GEO. This will give a good representation of what the network would have looked like if begun in 2006. This data shows that U.S. commercial satellites are the primary contributor to the network, with U.S. Military second. It also shows that using U.K. as a partner may not be worth the effort based on cost with only one communication satellite launched.

Name of Satellite, Alternate Names	Country of Operator/Owner	Class of Orbit
Echostar 14	USA	GEO
Intelsat 16 (IS-16)	USA	GEO
DirecTV-12	USA	GEO
Intelsat 15 (IS-15)	USA	GEO
Intelsat 14 (IS-14)	USA	GEO
TerraStar 1	USA	GEO
Sirius FM-5	USA	GEO
SES-7 (Protostar 2, Indostar 2)	USA	GEO
Wideband Global Satcom 2 (WGS-2, USA 204)	USA	GEO
Galaxy-19	USA	GEO
AMC-21 (Americom 21)	USA	GEO
Echostar 11	USA	GEO
Intelsat 25 (IS-25, Protostar 1, Chinasat 8)	USA	GEO
Galaxy-18	USA	GEO
ICO G1	USA	GEO
DirecTV-11	USA	GEO
SDS III-5	USA	Elliptical
Wideband Global Satcom 1	USA	GEO
Intelsat 11 (PAS 11)	USA	GEO
Spaceway 3	USA	GEO
DirecTV-10	USA	GEO
Galaxy-17	USA	GEO
AMC-18 (Americom 18)	USA	GEO
WildBlue 1	USA	GEO
XM Radio 4 (Blues)	USA	GEO
DirecTV-9S	USA	GEO
Galaxy-16	USA	GEO
Echostar 10	USA	GEO
GE-23 (AMC-23, Worldsat 3, Americom 23)	USA	GEO
Galaxy-15	USA	GEO
Galaxy-14	USA	GEO
Galaxy-28 (G-28, Intelsat IA-8, Telstar 8)	USA	GEO
DirecTV-8	USA	GEO
Spaceway F1	USA	GEO
XM Radio 3 (Rhythm)	USA	GEO
INMARSAT 4 F3	United Kingdom	GEO
Skynet 5C	United Kingdom	GEO

Skynet 5B	**United Kingdom**	**GEO**
Skynet 5A	**United Kingdom**	**GEO**
INMARSAT 4 F2	**United Kingdom**	**GEO**
INMARSAT 4 F1	**United Kingdom**	**GEO**
Thaicom-5	**Thailand**	**GEO**
Thaicom-4 (Ipstar 1)	**Thailand**	**GEO**
Superbird 7 (Superbird C2)	**Japan**	**GEO**
BSAT-3A	**Japan**	**GEO**
Amos 3	**Israel**	**GEO**
Nimiq 5	**Canada**	**GEO**
Telstar 11N	**Canada**	**GEO**
Nimiq 4	**Canada**	**GEO**
Anik F3	**Canada**	**GEO**
Anik F1R	**Canada**	**GEO**
Optus D3	**Australia**	**GEO**
Optus D2	**Australia**	**GEO**
Optus D1	**Australia**	**GEO**

Table 4. Hosted Payload Satellites

For the purposes of this study, actual communications equipment was not modeled. Instead, the core philosophy of limiting errors to the network and not the physical layer was held and all links were designed to build a physical network using the ephemeris of the chosen satellites. Also, some additional non-communications satellites were added so that a wider variety of orbital schemes were present, specifically non-communications satellites in low Earth orbits such as Worldview and Orbview imagery satellites.

The resulting network is much more highly randomized than any of the designed constellations and the interfaces between them tend to be much more dynamic, especially in the LEO satellites. The two ground stations used were Camp Roberts, CA and Singapore and each had a single antenna pointed toward the GEO belt (straight up for Singapore while Camp Roberts is pointed at an elevation of 60 degrees) with a 30 GHz carrier frequency and a beam width of 25 degrees. There was no specific antenna added for connection to the LEO satellites and so connections from ground stations to the LEO satellites can only occur when the satellite passes through the 25-degree cone of the antenna pointed at the GEO belt. This is displayed in Figure 33 and 34

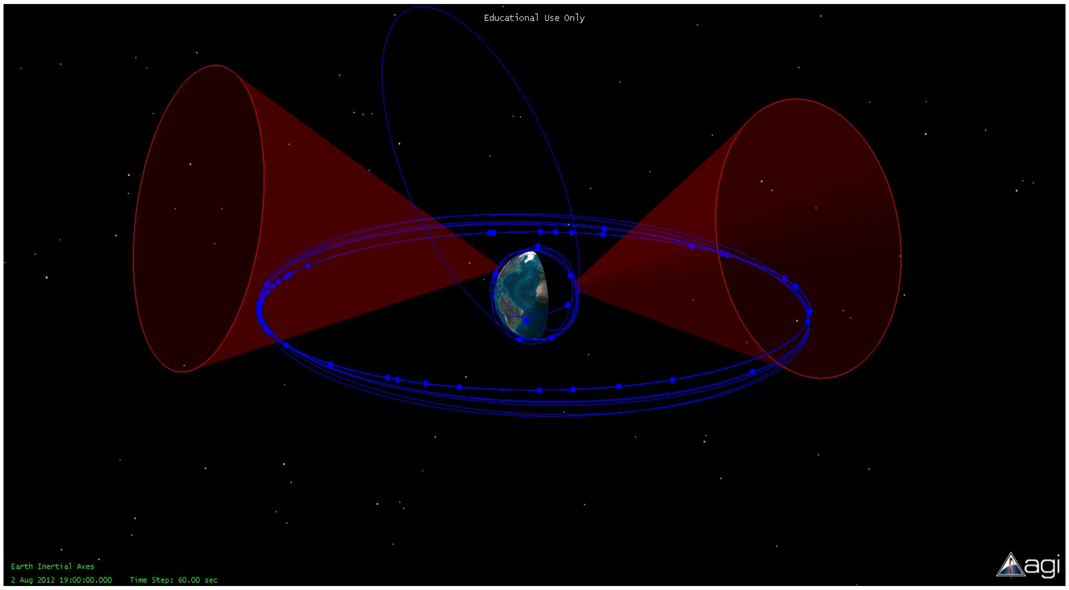

Figure 33.   Hosted Constellation 3D view

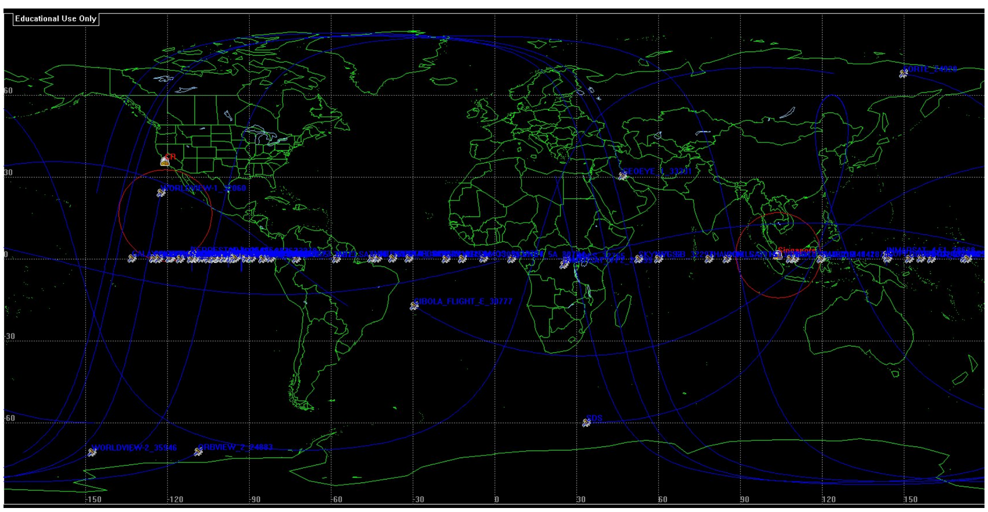

Figure 34.   Hosted Constellation 2D view

Each of the GEO belt satellites was given 3 antennas; a downlink antenna pointed directly at nadir with a beamwidth of 8.7 degrees (just enough to see the limb of the Earth) and two antennas for crosslinks. The crosslink antennas were pointed east and west with an elevation of 11.8 degrees and a beam width of 12 degrees. This way each crosslink was able to see several of the other GEO satellites to the east and west of itself. This is displayed in Figure 35

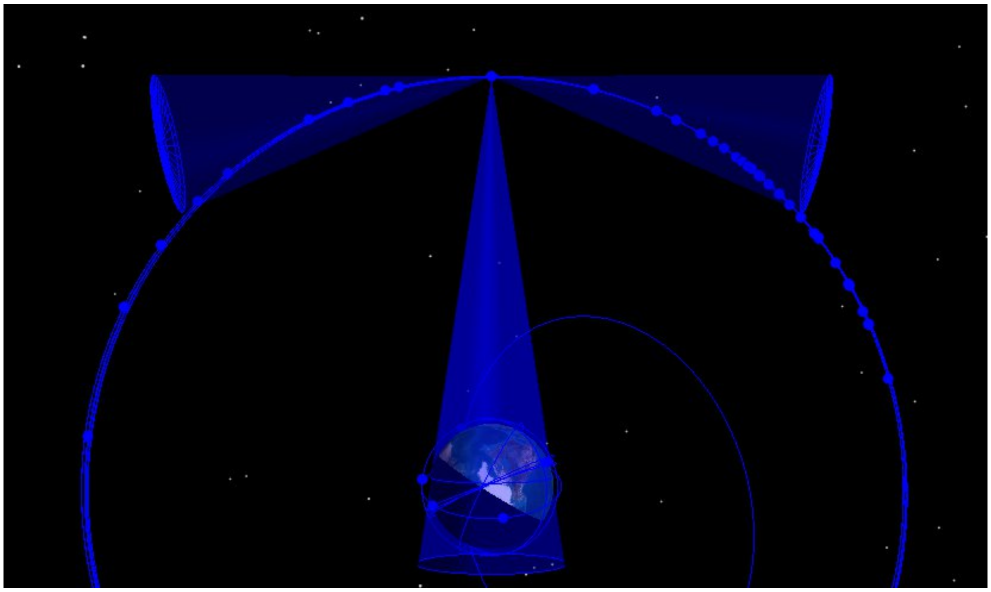

Figure 35. Hosted Constellation GEO Crosslinks

C. NETWORK TOPOLOGY MODELING

1. Wireless Networking

In order to create a network based on a constellation of orbiting nodes, it is important to distinguish a terrestrial comparison that can be initially followed for its design. There are two major types of networks that currently exist: fixed and mobile. The fixed Internet is comprised of all the PCs, routers, switches, and all other IP devices used for business, pleasure and academia. These nodes comprise a large and complicated network, but they do not change positions with any kind of frequency and their links are fairly static with respect to bandwidth and signal to noise aspects. Therefore, this is not a good analog to the satellite constellation based network. The second network type is the mobile network. This is often described as a Mobile Ad-hoc Network (MANET), which contains nodes of different types that are dynamically coming in and out of the network and change positions with time. This type of network is flat, with no physical infrastructure and the transmission medium is the RF spectrum. Every computer or device is considered a node and acts as a router and end host within the network topology. It must deal with limited computing resources,

limited energy and bandwidth, as well as a highly dynamic topology. For these reasons, the MANET network provides the best description for the Satellite Constellation Network. The following section describes the components of a MANET network.

2. OSI Layers

The basic model used to describe how nodes communicate within a network is the Open Systems Interconnection (OSI) model, or stack as it is often described. This standard divides network communications into seven layers: Physical, Data Link, Network, Transport, Session, Presentation, and Application (Dean, 2010). The OSI model is displayed in Table 5.

	Data	Layer	Routing
Host Layers	Data	**Application** Network Process to Application	
	Data	**Presentation** Data Presentation and Encryption	
	Data	**Session** Interhost Communication	
	Segment	**Transport** End-to-End Connections and Reliability	
Media Layers	Packet	**Network** Path Determination and IP (Logical Addressing)	
	Frame	**Data Link** MAC and LLC (Physical Addressing)	
	Bit	**Physical** Media, Signal, and Binary Transmission	

Table 5. OSI Model

This model is a representation of how two nodes communicate with respect to the different protocols affecting how data is transmitted. The problem being analyzed will focus on the Physical, Data Link, Network and Application layers of this model. The Physical layer describes how the network node manages the signaling to and from network connections. This includes the antennas and radio output parameters that will be used for the satellite networks. The next layer is the Data Link layer which describes how data is packaged into frames that can be sent over the Physical layer. In this case this will be the Media Access Control parameter. The third layer for our analysis is the Network Layer. This layer is responsible with making decisions regarding how data should be moved through the network with respect to routing decisions at a logical level. The final layer that applies to this analysis is the Application layer. This provides the interface between software applications so that they can communicate data to the network level both up and down the stack. These parameters must be understood and identified before the satellite constellations can be modeled and evaluated in a test environment.

3. Physical Layer

The Physical layer of the SCNMs is responsible for setting the data transmission rate monitoring data error rates, but is unable to correct for errors. In order to analyze the satellite SCNMs in a modeling environment the Physical layer variables were standardized based on results from the STK modeling. This was done to ensure that the satellite and ground station antennas would operate at an appropriate power level for connections to be made as long as there was a line of sight capability between nodes. While this may not always be the case in reality, this was not the focus of the research and was removed as a test variable. Treating the Physical layer components as constants among the nodes in the SCNM, the following values were used in the computer modeling environment:

Radio: Abstract Model

Data Rate: 125 Mbps

Transmitter Power: 60 dBm

Sensor Sensitivity: -58dBm

Threshold Sensitivity: -48 dBm

Reception Model: Signal to Noise Ratio Based, 10 dB

4. Data link layer

The second layer of the OSI model is comprised of the Logical Link Control and the Media Access Control. These sub layers work together to take data from the Network layer and create frames that can be transmitted over the Physical Layer. A Frame is a package for moving data that includes the raw data, but also sender/receiver addresses, error checking and control information (Dean, 2010). This layer is also responsible for ensuring an acknowledgement message is received from the receiver's Transport layer within a prescribed amount of time, before the transmission must be tried again. The Media Access Control (MAC) sub layer manages the access to the physical medium by appending the destinations nodes physical address or MAC address to the data frame being transmitted. There are a variety of different MAC protocols that are currently used for differing network types including TDMA, PSMA and CSMA. For the purpose of this modeling the SCNM was designed to use Carrier Sense Multiple Access with Collision Avoidance MAC protocol. Using this protocol the node senses the channel before it sends data. If the channel is idle, data is immediately sent. If the channel is busy, a backoff timer starts. With the collision avoidance policy set, this backoff timer is paused when the channel is busy and resumes when the channel becomes idle (Qualnet 2011).

5. Network layer

a. Introduction

In order for information to travel across a computer network there are often a number of paths that the data can travel and depends on the type of network. A wired network with fixed notes will have set routes that data can take, whereas a wireless network with mobile nodes complicates this problem. The process of identifying the appropriate path that data should be sent through a network is handled by the routing protocol. This routing protocol collects data about the status of the network and selects the best path for packets to reach their destination. A satellite network will be based on a mobile adhoc network with frequently changing node position.

A routing protocol for a MANET should be capable of handling a very large number of hosts with limited bandwidth and energy (Huhtonen 2004). What is unique to a MANET is that the protocol must also deal with host mobility and physical links constantly being created and broken as these nodal positions change with time. For this reason a MANET requires all hosts to generally act as a router, participating in route discovery and maintenance. Because a MANET network is normally composed of RF links between the nodes, the protocol must also reduce the overhead required to discover and maintain these routes due to the limited bandwidth available as compared to a traditional wired network. In order to overcome these challenges, network engineers have created two major types of routing protocols: Proactive and Reactive.

b. Proactive Routing

The first main category of MANET routing protocols is Proactive in nature. These protocols are table-driven and will attempt to constantly maintain up to date routing information or tables. This is done by sending control messages or queries for updates from other nodes within the network. Whenever a change is detected in the status of a node or link, this information is propagated throughout the network (Huhtonen 2004). There are two main subclasses to

Proactive routing protocols: event-driven and regular updated. Event-driven protocols will only sending routing update packets if a change in the network topology is detected. Usually this would be because a link is broken or a node enters or leaves the detectable area. In this case, the node that detects the change would forward this information based on the protocol design. The second subclass is regular updated protocols. These protocols will send their topology and link state information at regular intervals. Often times this will be designed so that closer nodes update more frequently than further ones, in order to balance the load on the network from this overhead (Lang 2003). An advantage to this type of scenario is that the overhead produced by these discovery and update messages is predictable and relatively static. A disadvantage is that the overhead may be rather large with most of the data unnecessary with respect to routes and links that are not necessary or required at any given time. The Proactive method gives a global network view, but is slow to converge or adapt to topology changes.

c. Reactive

A reactive or on-demand routing protocol is designed such that routes are created only when needed by the host and only maintained for that period. In this case, the routing information is polled from the destinations, as opposed to being pulled as would be the case for a proactive protocol. When a node has traffic to send, it will broadcast a message for route information to its nearby neighbors. This is usually done to a nearby set of neighbors, in hopes they will have the route information required. If this doesn't work, then the broadcast will be expanded. The goal is to minimize the effort in polling for the route and thereby reduce the overhead in the network. Route information can be cached for common routes, but will expire with time or as the mobile network topology changes. Another issue is the initial setup delay as the route is being discovered before the message can be send, which is not required in the proactive routing scheme. The advantage of this type of routing protocol is that RF bandwidth is not being used to transmit topology information for routes that

are not being utilized (Lang 2003). The Reactive method only gives a partial network view, but is quicker to react to network topology changes.

d. Hybrid

Some routing protocols seek to combine the Proactive and Reactive methods in order maximize bandwidth and reduce overall delay. Most of these protocol types are not a true hybrid. Rather they use proactive methodology for a local region where node position does not change greatly, and a reactive methodology outside of this local zone. In theory, this type of protocol should give the best solution, but has not been as thoroughly explored as the primary types due to the complexity involved. An example of this is the Enhanced Interior Gateway Protocol (EIGRP) designed by CISCO.

e. Route Selection

There are generally five major strategies that can be used for the actual selection of routes. The first two are Signal Strength and Link Stability. These tend to be related to each other as they both involve the physical layer. In the case of signal strength decision making, packets are routed along the connection with the best signal strength. In the case of Link Stability, packets are routed along the connections that appear most stable over a period of time. This could be due to physical effects of the RF connection itself such as atmospheric losses or due to the variation in nodal position with mobile nodes. Another variation is to judge the routing decision on a specified metric such as the shortest path, hop count, or the link state. In the case of hop count, the protocol is often described as a distance vector protocol. A protocol of this type may route based on geographical distance, shortest number of hops, or even move packets in the direction of the destination without focusing on a specific route. Finally, the link state protocols look to the specifics of the link itself to make routing decisions. This may be based on quality of the link, traffic load or other factors (Lang 2003).

6. Routing Protocols in Experiment

a. Experimental Choices

A total of 5 routing protocols were chosen for testing across the SCNMs. These protocols were chosen in order to judge how the different route decision methods, network overhead, network size, and mobility would change across different protocol types. To give a sampling across different routing styles, the team chose two reactive, two proactive and one baseline protocol. An introduction and general explanation for each of these protocol types will be given in the following sections.

b. Ad-hoc On-Demand Distance Vector

Ad-hoc On-Demand Distance Vector (AODV) is a Reactive or on-demand routing protocol designed specifically for use in mobile ad-hoc networks operating in either IPv4 or v6 at the network layer (Qualnet 2011). Because this is a reactive protocol, routes are created and maintained only as needed. To maintain the routing tables, the protocol stores information regarding the next hop to the destination and a sequence number indicating the freshness of the data. Information relating to the active neighbor nodes is received through the discovery of the destination node. As a corresponding route breaks, this information is then passed to the neighbor nodes.

The process for sending data via AODV begins with a Route Request (RREQ) message being broadcast to neighboring nodes with the requested destination sequence number. This sequence number is used to prevent old information being returned and minimize looping problems. As this RREQ message continues through the network to the destination, only the sequence number or hop metric is increased, but other data is not added about the passed hosts. These intermediate nodes will update their routing tables with respect to the requested host. This allows the destination node to reply easily to the initiating host as it travels back through the network. Finally a Route Reply (RREP) message is sent back to the initiating host with the final route to send

data. This can come from either the destination host or an intermediate host who has the information that the destination host is active and the connection is fresh. To maintain the route, an intermediate host will send a Route Error (RERR) message to the neighbor nodes if a link is broken. Hello messages are also periodically sent for maintaining current route information.

Older distance vector protocols were prone to looping issues, but this is solved here through the use of the sequence numbers. Also, this protocol seeks to reduce the control traffic overhead but incurs a latency increase in the discovery of new routes. In previous testing, AODV has been found to react relatively quickly to topological changes in the network and update only the nodes affected which further reduces control overhead (Huhtonen 2004).

c. Dynamic MANET On-Demand

The second reactive protocol chosen for comparison is Dynamic MANET On-Demand (DYMO). It is a unicast reactive routing protocol for use by mobile nodes in a wireless multi-hop network (Qualnet, 2011). DYMO operates in a similar way to AODV, broadcasting RREQ packets to discover network when a message will be sent. The difference between the two is that DYMO will discover the routing information for all nodes in the path to the destination. In this case, all nodes between source and destination will exchange routing information. This reduces overhead when resending packets over routes that have already been discovered. In a network with highly mobile nodes, this can become a disadvantage as links are more frequently broken, causing packets to be dropped due to no route. When this happens the overhead caused by sharing routing information of all nodes in path ends up being higher than with AODV (Chung et al. 2008).

d. Optimized Link State Routing

The Optimized Link State Routing (OSLR) protocol was developed by the French National Institute for Research in Computer Science and Control (INRIA). It is a table-driven, proactive protocol that updates routing information on

a periodic basis. This protocol seeks to minimize control traffic by designating a small number of nodes as Multi Point Relays (MPRs) for flooding topographical information (Qualnet 2011). These nodes act as a repository for network topology information and are accessed in order to form an optimal route for a node to its destination. This reduces flooding of broadcasts by minimizing the same broadcast in a region.

The OLSR protocol uses two main types of messages for control. The first is a Hello message which is used to discover status of local links and neighbors. These messages are only forwarded one hop away. It is used to determine if the neighbor is active and whether the link is unidirectional or bidirectional. The second message is a Topology Control (TC) message. TC messages are broadcast over the entire network and include information about advertised neighbors and MPR lists. The unique concept behind OLSR is the MPRs and how they reduce information exchange overhead. The MPR is the one hop neighbor for a node and can forward its messages. For any local group of nodes there is only one MPR, so the control overhead is minimized (Huhtonen 2004).

As a proactive protocol, OLSR has the advantage of always having routing information for all nodes within the network. With this method, the protocol will periodically send updated topological information across the network. The frequency of this periodic update can be adjusted based on how mobile the network is, but more frequent updates result in a higher network overhead in control messages. This protocol is well suited to a dense network where there is minimal physical delay in network transmission and the nodes can communicate and share topographical information more easily Huhtonen, 2004).

e. Source Tree Adaptive Routing

Source Tree Adaptive Routing (STAR) is a proactive, link-state type routing protocol that utilizes a source tree approach to calculating routes within the network (Qualnet 2011). The source tree is computed from information about

a node's links and neighbor source tree information which is then formulated into a list of preferred links to all other destinations. The source tree information is then shared with neighbors periodically. Each node will use the source tree information to compute a routing table with destination and next hop for an outgoing message. This protocol uses Link State Updates to share source tree data with neighbor nodes. There are two modes in which STAR can operate. The first is a least overhead routing approach and the second is an optimized routing approach for shortest path (Qualnet 2011).

To conserve bandwidth, STAR is designed to only communicate changes to its source routing tree when a router detects specific changes. These can include new destinations, looping, node failures or network partitions. The least overhead routing method of this protocol attempts to approach the results a reactive protocol may attain while still constantly maintaining all routes (Ekberg 2008).

f. Open Shortest First Path Version 2

Open Shortest First Path (OSPFv2) is proactive, link-state routing protocol designed for use in fixed terrestrial networks. It is an Interior Gateway Protocol that calculates its routing tables based on a link-status algorithm which follows Djikstra's shortest path method (Comer 2004). It accomplishes this by periodically probing adjacent routers, followed by broadcasting a link status message. In OSPF each router in the network will receive the broadcast message, update its local routing table and recompute shortest paths. This protocol is designed for use in an autonomous system, not a MANET, such as a specific provider network or Internet. It was not designed for use in a dynamic environment with mobile nodes and constantly changing links. It was chosen for use as a baseline protocol for comparing the performance of the MANET protocols.

7. Application Layer

The Application Layer resides at the top of the OSI stack. It is designed to allow software application to interpret data that is coming up from the network layer and to communicate data down as well. This is how a software program can negotiate things such as formatting, security, synchronization and other tasks with the network (Comer 2004). In the case of this analysis, the Application Layer will not be manipulated once set within the test environment. Rather, standard applications will be determined and emulated over each test regime as a constant. These simulated Application Layer programs will provide the traffic generation in order to analyze the network behavior.

D. NETWORK MODELING

1. Qualnet Introduction

After completing the design of the constellations, the next phase of the analysis was to model the network behavior over them. In order to do this there were a number of network and modeling simulation software programs considered that exist for the general purpose of network emulation and evaluation. Three primary candidates were selected for this purpose and evaluated based on the needs and resources of the research team. These networking programs were iTrinegy, Opnet and Qualnet. The iTrinegy emulation software is capable of modeling satellite network operations but was not selected due to cost. Opnet is a popular modeling software, but is not well suited to working with satellite based networks and is not very flexible when it comes this type of simulation. Qualnet was decided to be the best choice for a variety of reasons that will be explained further.

Qualnet is a suite of tools produced by Scalable Network Technologies Inc. It provides a comprehensive environment for modeling wired and wireless networks, designing protocols and analyzing their performance. The software suite also has a vast library of models that can be added in order to use currently defined standards, protocols and vendor’s products for simulation. These include

Developer, Wireless, Multimedia, Enterprise Model Libraries, sensor networks, satellite, and cellular models (Qualnet 2011). In addition to this, Qualnet has worked in concert with AGI in order to design a Qualnet plug-in for use within STK. This allows a scenario to be created within STK to be modeled by Qualnet through a GUI without having to leave the STK program. In order to accomplish this STK models the dynamic positions of assets, considers antenna pointing and computes link budget metrics. This data is imported into Qualnet through the plug-in where it can then be used to analyze network performance. Due to the complicated nature of the constellations being considered in this analysis, this was chosen as the best method for modeling the routing protocols and network behavior.

2. Setup

After each of the four constellation types was fully modeled in STK, they were then imported into Qualnet. At this point each SCNM underwent a four step process to set all the networking variables prior to testing the protocols. The following sections will describe this process and identify the variables that were set. Due to the extremely large number of setting that can be adjusted within the Qualnet interface, all variables were kept at default values unless described below.

a. Scenario Configuration

The first step was to set the Scenario Configuration. This consists of adjusting global settings that will apply to all of the imported nodes from STK. In this section the frequency of the channels is set. For the purpose of this simulation it was decided to model each link as a single bi-directional channel capable of full duplex at 30Ghz (Ka band). This was decided based on the trend towards Ka band for future communication satellites in production due to the increased bandwidth and data transfer capability. Figure 36 displays the setup of

Scenario Configuration parameters in the Qualnet GUI. For all screenshots displayed from Qualnet GUI, fields in bold font are those changed from default values.

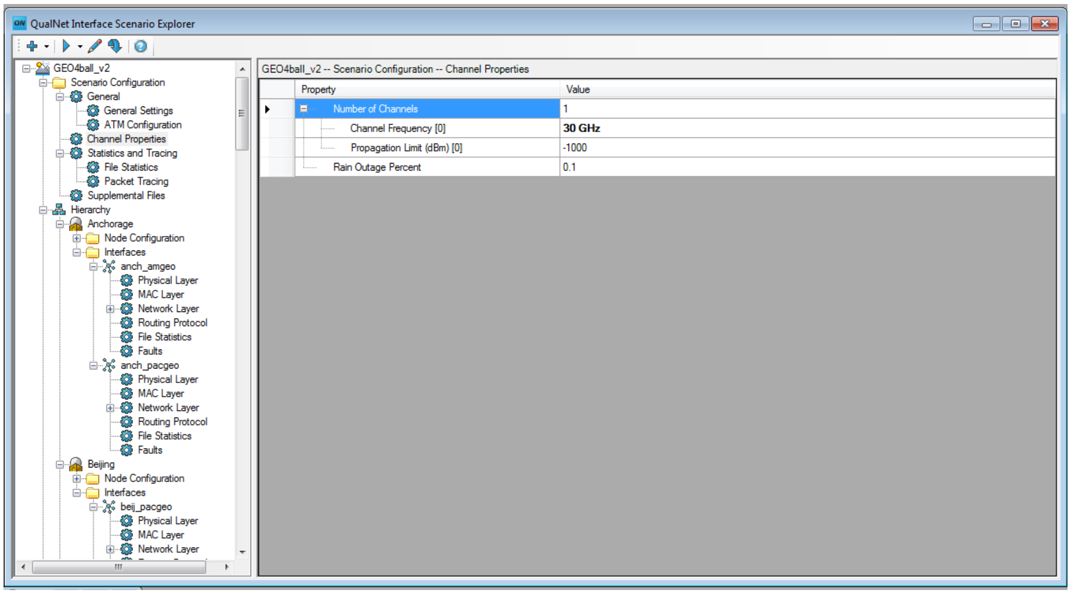

Figure 36. Scenario Configuration

b. Hierarchy

The second step in the setup process is to build out the Hierarchy section. When the GUI is first opened and the assets are imported into Qualnet, they are grouped into a list under the Hierarchy section with each asset or node listed as they would be in STK. Under each node is a folder for the node settings and interfaces. At this point, interfaces must be created for every link that exists in the network between the nodes. It should be noted that the OSI layer networking variables can be set at 4 different levels within the modeling tool including global, node, interface or subnetwork. According to Qualnet personnel, settings made at the subnet level will be carried over to the node and interface levels, but in experimentation this is not the case. After reviewing Qualnet documentation it was found that the precedence rules for OSI variables are Interface, Subnet, Node and Global with preceding settings applying to the other layers.

The process for building the Interfaces is similar to adding a network card for each link that exists on the node. In the case of the GEO SCNM, the ground nodes each have two interface and the GEO satellites each have four interfaces. These numbers increase drastically for the LEO and Hybrid SCNMs and are summarized in Table 6 After the Interfaces are created, each must have the Physical, MAC, Network and Protocol settings input. This is done for each node in the hierarchy. The settings used for the Physical, MAC and Network layers are displayed in the Figures 37 through 40.

Model	**Number of Nodes**	**Number of Interfaces**
GEO	8	24
LEO	68	331
Hybrid	72	413
Hosted	65	181

Table 6.   Interface Count by Model

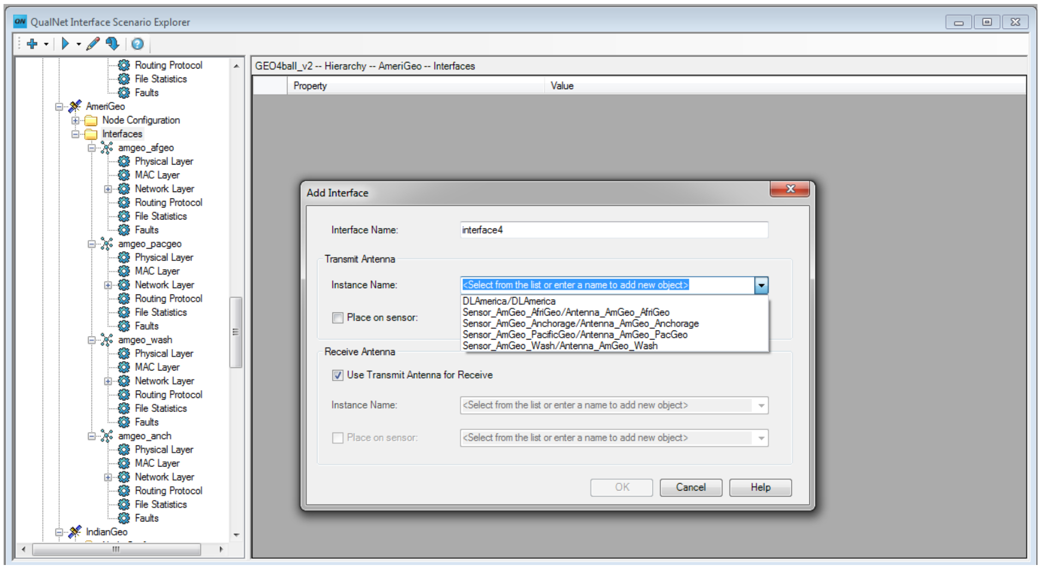

Figure 37.   Interface Creation

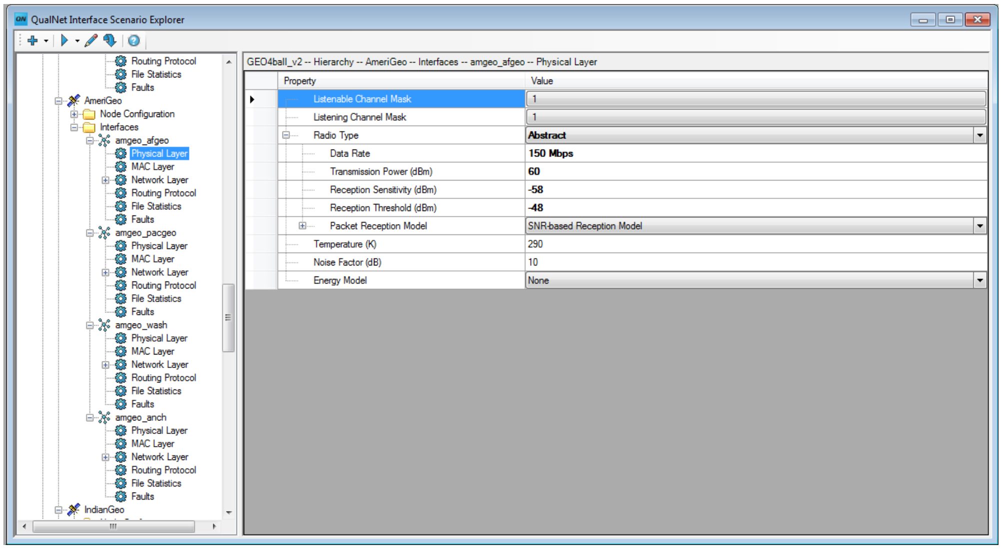

Figure 38.   Physical Layer Settings

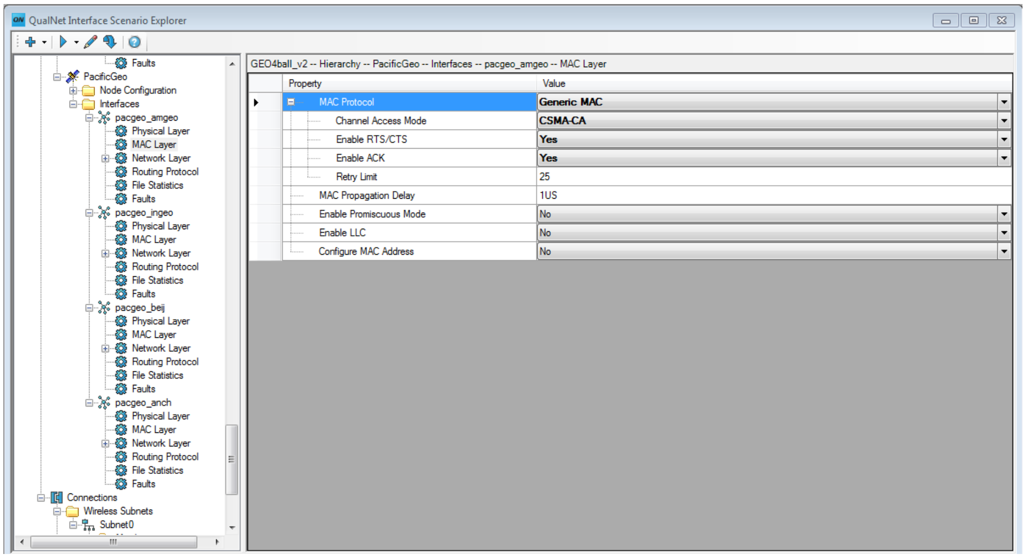

Figure 39. MAC Layer Settings

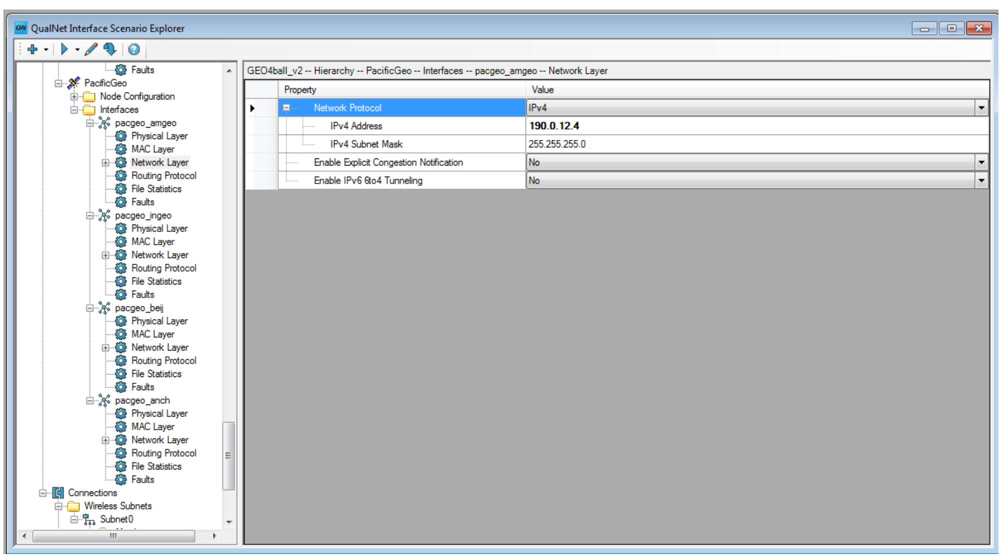

Figure 40. Network Layer Settings

c. *Connections*

After building all of the interfaces and inputting the appropriate OSI settings for each SCNM, the next step was to add all of the interfaces into a

subnet. By doing this it allows the nodes and interfaces to connect with each other on one network. In the case of the GEO and Hosted models this was as simple as adding a subnet, selecting all the interfaces, and adding them to the subnet. At this point the same OSI setting from above are input at the Subnet level. In the case of LEO and Hybrid SCNMs a problem occurs during the automatic process creating the subnet due to the volume of interfaces. At the network level a subnet with a subnet mask of 255.255.255.0 can have 255 interfaces added with distinct IP addresses in IPv4. Because the other two models have more interfaces than that, the subnet mask had to be changed to 255.255.0.0 and IPs had to be manual input into each of the interfaces.

The final step of the setup before the network simulations are run is to create applications that can run over the network. These applications are basic traffic generating protocols developed by Qualnet that enable different networking metrics to be recorded based on the type of program. In order to gather a variety of metrics about the networks and test how different types of traffic affected them, four traffic generators were chosen and added under the Connections. These were Constant Bit Rate Client, Variable Bit Rate Client, Lookup and Super-Application. These traffic generators are summarized in Table 7.

Traffic Generator	Description
Constant Bit Rate (CBR)	This UDP-based client-server application sends data from a client to a server at a constant bit rate.
Variable Bit Rate (VBR)	This model generates fixed-size data packets transmitted using UDP at exponentially distributed time intervals.
Lookup	This is an abstract model of unreliable query/response traffic, such as DNS look-up, or pinging.
Super-Application	This model can simulate both TCP and UDP flows as well as two-way (request-response type) UDP sessions.

Table 7. Traffic Generator Applications (from Qualnet 2011)

3. Initial Testing

In order to test the functionality of the Qualnet STK plugin, two initial scenarios were built and run through STK. The first was a familiarization tutorial from the STK Help Menu. This tutorial involves adding a UAV and vehicles ground vehicles to STK along with associated antennas. The scenario is such that the ground vehicles are connected via 802.11b network and one of the vehicles is capable of communicating with the UAV through a tracking UHF antenna. This gave a basic familiarization with the functions and setup of the Qualnet GUI.

The second test scenario was built involving one LEO satellite and one ground station. The satellite was set at Camp Roberts, CA and the satellite was an Iridium satellite. The time frame was a 15 hour period such that the satellite started in view of the ground station, broke contact for approximately 12 hours and then regained contact for a second pass. This scenario was designed in order to determine if the Qulanet Plugin was properly importing the mobility statistics of the satellite nodes as they move through their orbit. More specifically, it showed that the program properly used the link budget data from STK to

determine when the ground station and satellite were in reception range and adjusted the packet transfer in the network accordingly.

4. Process

Before an SCNM could be simulated, a final step requires setting the Routing Protocol at every Interface and Subnet level. As previously discussed five protocols were chosen for comparison: AODV, DYMO, OLSR, STAR and OSPFv2. Each of these routing protocols was simulated for the four SCNMs: GEO, LEO, Hybrid and Hosted, providing twenty total data sets. Each run takes approximately two hours for the computer to simulate the traffic applications discussed for a period of 100 seconds. After each simulation, the data is stored by Qualnet in a .stat file for later use. From these data sets, the following metrics were recovered for use in comparing the sixteen scenarios: IPinRecieves, IPoutRequests, IPinDelivers TTL-Based Average Hop Count, Data Packets Dropped for No Route and Longest Time in Queue. These metrics are defined by Qualnet in Table 8. A variety of other data is also available but not used for this analysis. Particularly, Physical Layer data was not used because the settings for this layer were designed such that connection was always possible when two nodes' antennae were within line of site. The reason for this was the goal of testing the network protocol compatibility with the various constellations of mobile satellite routers, rather than the effect changing link conditions had on the routing protocols. Figure 41 is an example of the output data in raw form from the Qualnet Plugin. The metrics of interest were transferred from this format to excel for further analysis.

Qualnet Output Metric	Description
IPinRecieves	Packets received from the MAC protocol.
IPoutRequests	Packets sent by IP to the MAC protocol.
IPoutNoRoutes	Packets dropped due to unreachable destination.
IPinDelivers TTL-Based Average Hop Count	TTL based on average hop count for packets sent to upper layers.
Longest Time in Queue	Longest time spent in queue by a packet (in seconds)

Table 8. Qualnet Output Metrics (from Qualnet 2011)

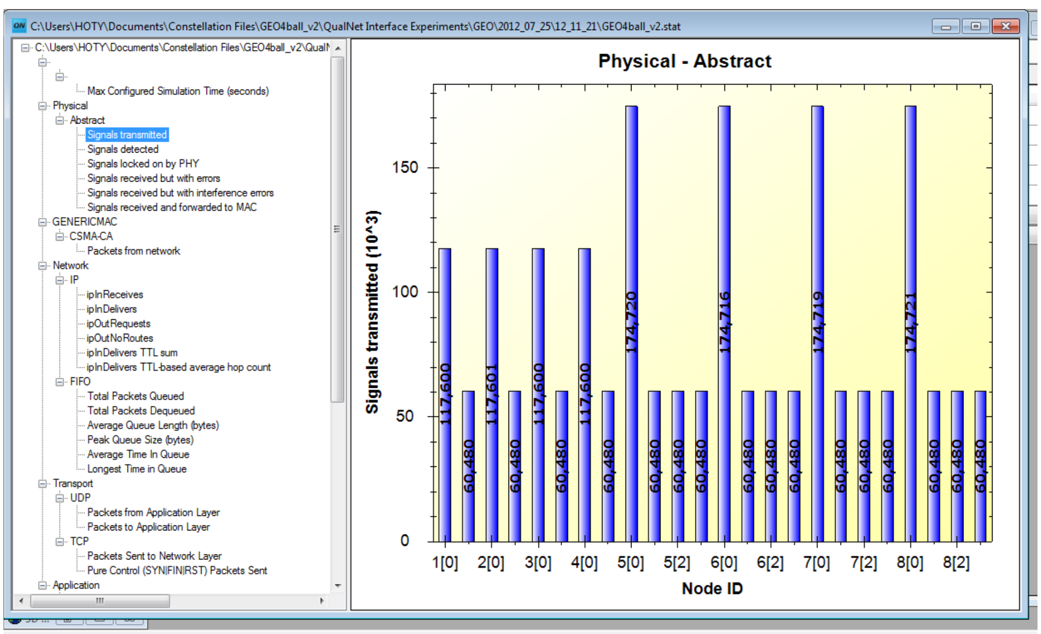

Figure 41. Results from Qualnet Simulation

5. Shortcomings

A number of shortcomings in the STK Qualnet Plugin were discovered over the course of the experimentation. The first issue is the propagation of

settings based on the current precedence rules. In a network composed of a small number of nodes, this is not a major issue. In the case of the Hybrid model the requirement to set the OSI variables 413 times is a major oversight by programmers. This would easily be remedied by allowing the subnet settings to propagate up to the node and interface settings with the same result. For each simulation run, it would take approximately an hour to change each interface to the new protocol, when it would only take one setting change at the subnet level.

Secondly, this plugin does not have a feature to allow data to be exported for analysis. The data display and results window is good for an initial cursory look at the data, but once again is not well suited to large numbers of links or nodes. Creating a tool for exporting the data to Excel or another analysis tool would add needed functionality.

Furthermore, the data is only in an aggregated total and there is no way to view how the rates changed with regards to time. If there were a way to select a specific node or interface and view the changing metrics with regards to time in order to diagnose networking problems it would be very helpful. STK relates nearly every piece of data back to a specific time, the same should hold true for the Qualnet plugin.

With regards to the channel setup in the Qualnet Plugin, it would be more useful to have the frequency directly imported from STK rather than manually input. This way a network using interfaces with multiple frequencies would be easier to model.

Finally, the STK Qualnet Plugin suffers from a lack of documentation. The only dedicated reference is the Help menu which contains a few short pages. Compared to the full Qualnet Software which has a very robust set of User Guides, this STK interface is extremely lacking. Particularly since the Plugin has a number of features that are different from the standard Qualnet GUI, the lack of documentation makes the software less useable than it could be.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. Modeling Results and Analysis

The modeling effort generated 16 sets of data; 4 constellations each with data for four different protocols (the STAR protocol failed to produce any data). Each set of data was then analyzed across each of the variables discussed in the previous section. This analysis focused on making relative comparisons between protocols within constellations and constellations within protocols.

A. CONGESTION

The total IPinReceives across an entire constellation as well as the total IPinReceives divided by the number of nodes was used to give a general idea of network Congestion. The basis for this as a corollary to congestion is that the number of actual data transfer packets being queued by the traffic generators in QualNet is a fixed number that does not change between protocols or between constellations, therefore any differences in total number of packets transferred must be due to network overhead vice traffic. This is displayed in Figures 42 and 43 on a logarithmic scale.

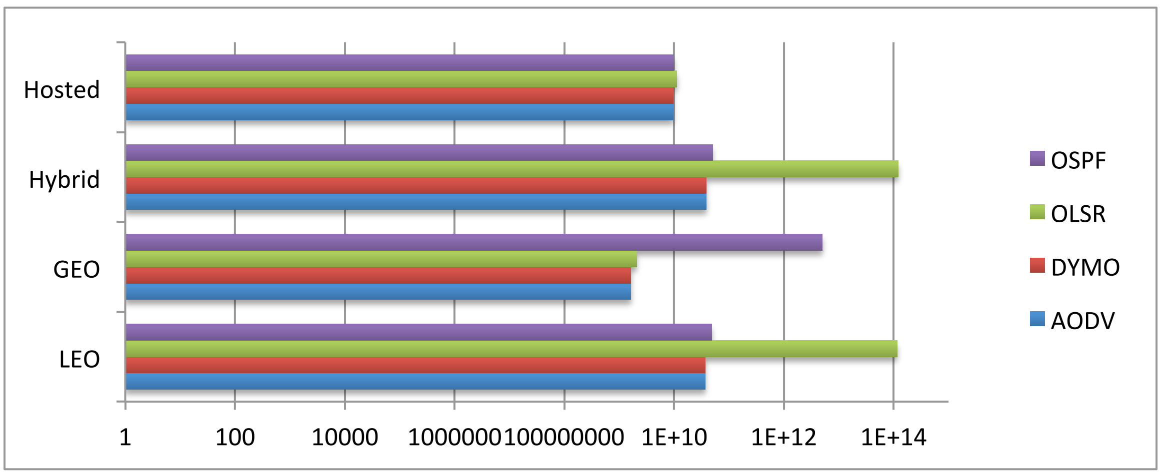

Figure 42. Total IPinReceives by Constellation and Protocol (log scale)

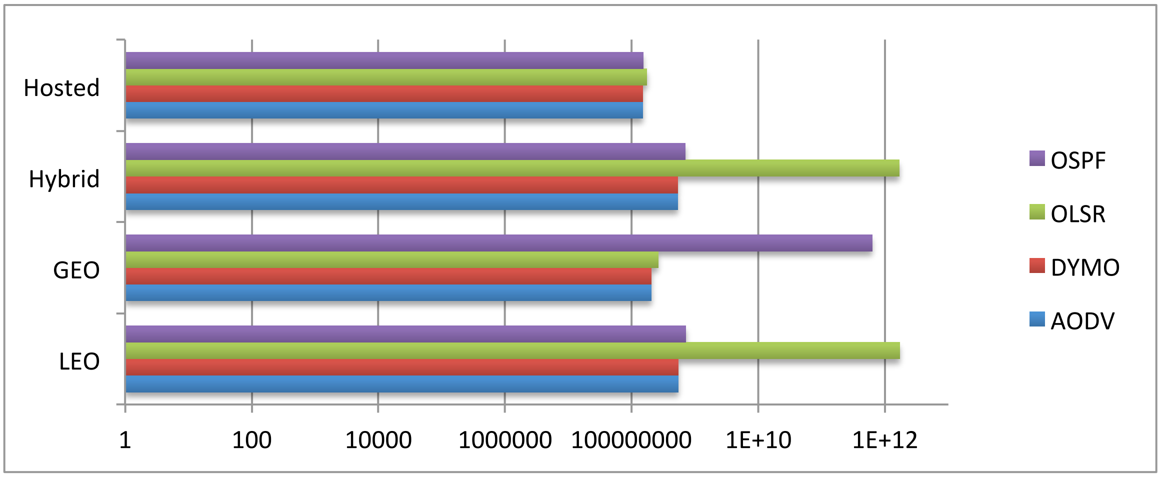

Figure 43. IPinReceives per Node by Constellation and Protocol (log scale)

In terms of total packets, the GEO constellation has the lowest congestions, as is to be expected since the GEO constellation has many fewer individual nodes flooding the network with discovery messages. Also of note is that the OLSR protocol had the highest congestion in all but the GEO constellation, where OSPF was the highest. In terms of total packets per node, the Hosted Constellation was the least congested. This is likely due to the high number of GEO satellites which are very static and therefore the network discovery can be achieved quickly and changes little over time.

B. NETWORK EFFICIENCY

In order to measure network efficiency in relative terms, each node was evaluated by dividing the total IP packets sent by the total IP packets received, and this number was further divided by the number of interfaces on that node to eliminate the skewing of the data due to multicasting. If data routing effectiveness is the external measure of a policy's performance, efficiency is the internal measure of its effectiveness. To achieve a given level of data routing performance, two different policies can expend differing amounts of overhead, depending on their internal efficiency. Protocol efficiency may or may not directly affect data routing performance. If control and data traffic must share the same channel, and the channel's capacity is limited, then excessive control traffic often

impacts data routing performance (Corson and Macker 1999). In satellite networks where the overall bandwidth available is very limited, the ability of a protocol to achieve data transfer with a minimal amount of overhead is paramount. The best way to measure the efficient use of bandwidth using the data generated for this study is to take the ratio of packets received to packets transmitted.

In addition to the efficiency per node, the constellations and protocols are compared by the variance between nodes. The variance is an indication of the "tune-ability" of the protocol or constellation. For example, if a protocol shows a high level of variance in efficiency between nodes, that indicates that there are many factors affecting the number of packets being generated by each node and thus more capability to optimize the protocol for that network. Conversely, the lack of variation in a protocol indicates that the efficiency of each node is independent of the specific conditions of that node and therefore there is little that can be done within that network to optimize the performance of each node. In conclusion, the variance metric is a corollary for the level of dependence between network configuration and protocol efficiency.

1. Constellations by Protocol

The GEO constellation had the best efficiency using the OSPF protocol, also given the number of nodes in the GEO constellation it is easy to differentiate performance among the nodes. For example, the most efficient nodes in the GEO constellation are the transmitting nodes; they have a far higher packet out count that packet in count. The receive nodes have the lowest because they have a substantially higher packet receive count than a packet sent count and the four satellite nodes all have similar efficiencies. The same trend is repeated among the other protocols with OSPF being the most efficient, AODV and DYMO being the same in between and OLSR having the least efficiency. Figure 44 displays nodal efficiency by protocol in the GEO constellation, Figure 45 focuses in on the OLSR protocol.

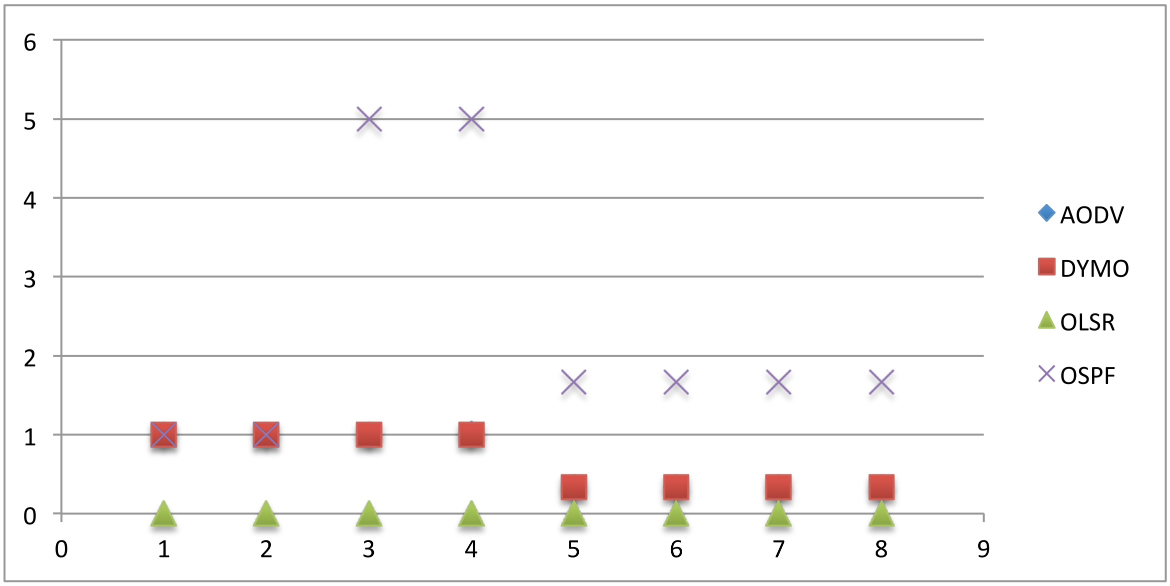

Figure 44. Node Efficiency by Protocol in the GEO Constellation (AODV and DYMO are identical)

By zooming in on the x-axis, we can see that the same pattern is repeated, at a much smaller scale, with the OLSR protocol.

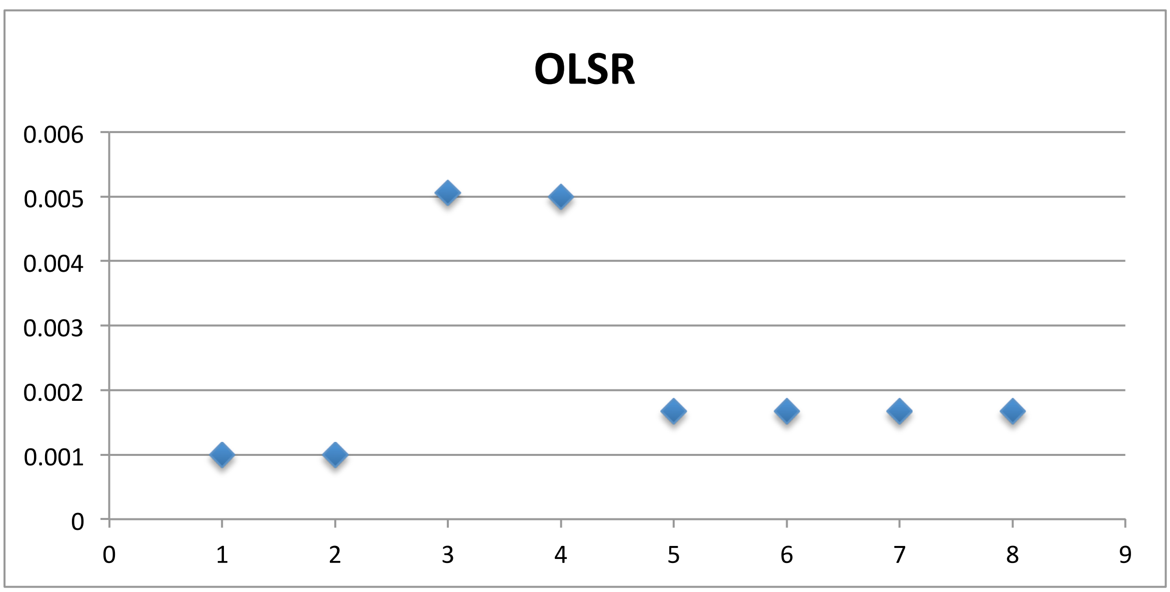

Figure 45. Node Efficiency of OLSR Protocol in the GEO Constellation

The variance between the different protocols within the GEO constellation also shows that the variation amung nodes using the OLSR protocol was the least and the variation of the OSPF protocol was the highest. This indicates that the OSPF protocol efficiency is much more dependant on network topology than the efficiency of the OLSR protocol within this constellation. This is displayed in Figure 46.

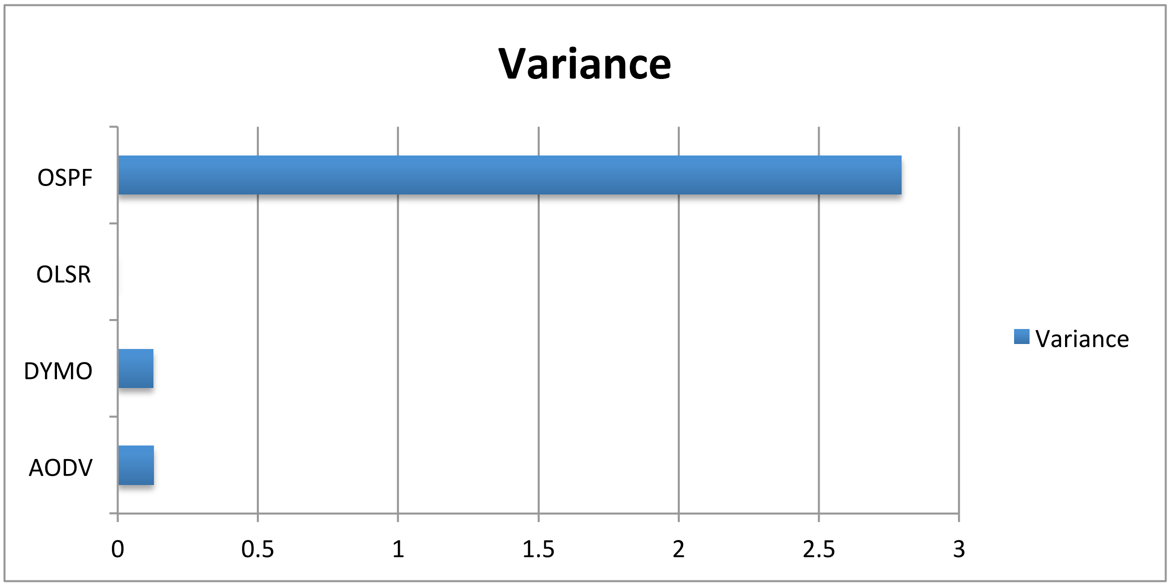

Figure 46. Variance of Nodal Efficiency in the GEO Constellation

The LEO constellation had good efficiency using the OSPF protocol, but all of the other protocols were substantially lower in efficiency; low enough to appear as a flat line with no distinguishing features. The result of OSPF being this efficient was unexpected as this protocol is not designed for mobile networks. This can be explained due to the fact that the simulation was only able to be run for a period of 100 seconds, in which time the LEO constellation topology changes very little. It is believed that this efficiency would go down the longer the simulation was run. This is displayed in Figure 47.

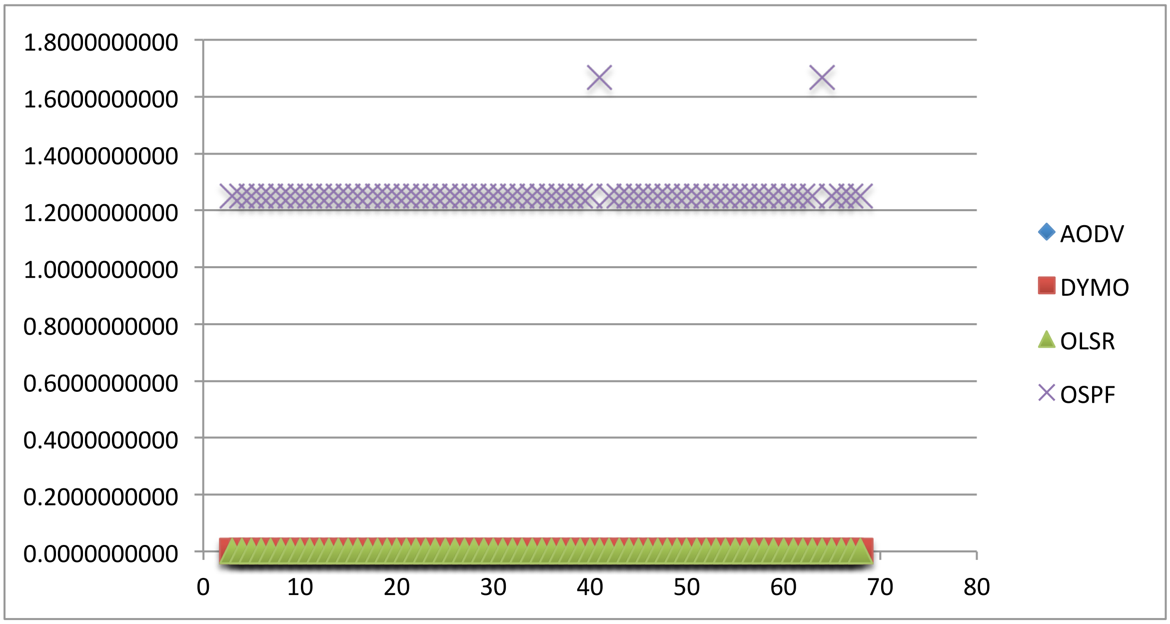

Figure 47. Efficiency of LEO Constellation by Protocol

By zooming in on the three protocols which appear to be on the x axis, we can see that while the two reactive protocols (AODV and DYMO) are very similar and with very little deviation among the various nodes, the OLSR protocol performed relatively better than the others. Efficiency of protocols in the LEO constellation is captured in Figure 48, while Figure 49 focuses on the OSPF protocol efficiency.

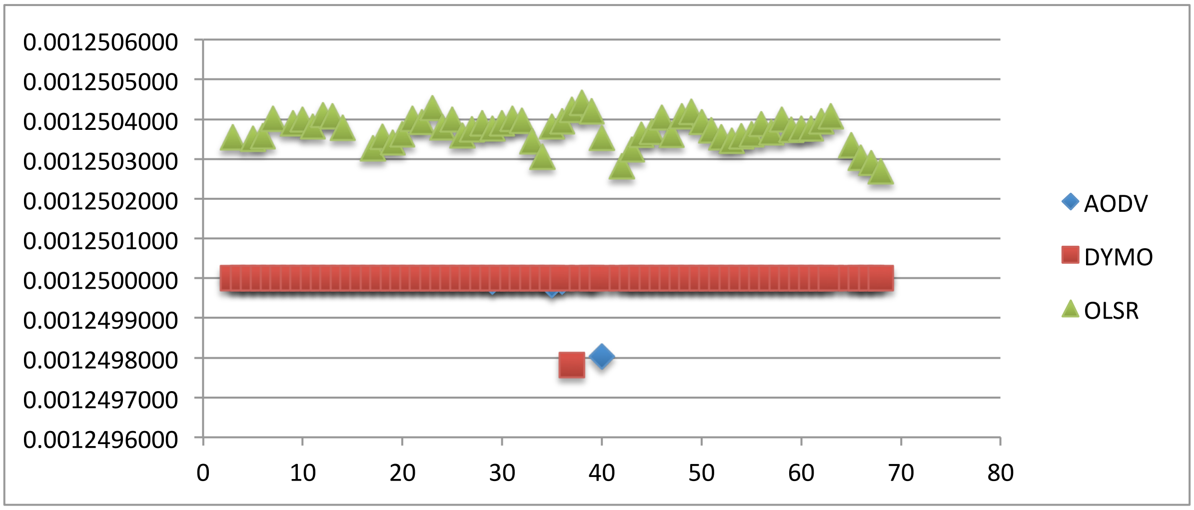

Figure 48. Efficiency of LEO Constellation by Protocols v2

Also, by zooming in on the OSPF numbers it can be seen that the OSPF efficiencies are similarly distributed to the OLSR efficiency, only at a much higher level.

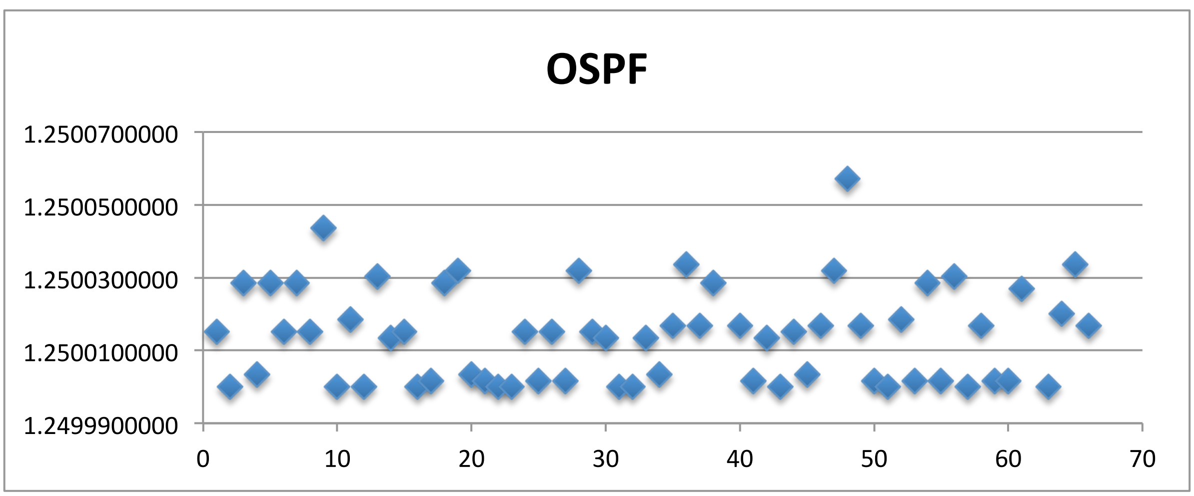

Figure 49. Nodal Efficiency of OSPF in the LEO Constellation

In addition, the calculated variance between the efficiencies of each node was analyzed to give an indication of how much the efficiency changed from

node to node. In the LEO constellation, the variance was very high for OSPF, and very low among the rest with OLSR having the lowest. The variance for the reactive protocols is practically zero but because of a few outliers in the data, the calculated variance is much higher. Figures 50 and 51 captures the variance of nodal efficiencies by protocol for the LEO constellation.

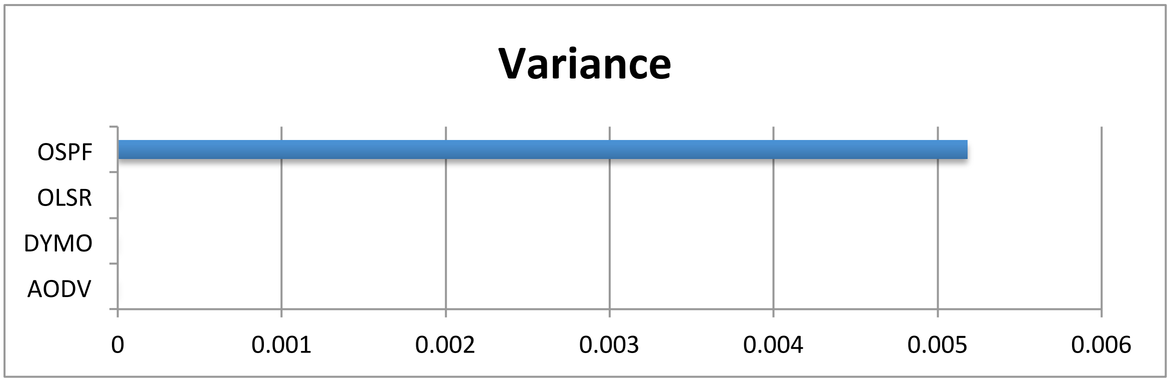

Figure 50.   Variance of Nodal Efficiencies by Protocol for LEO

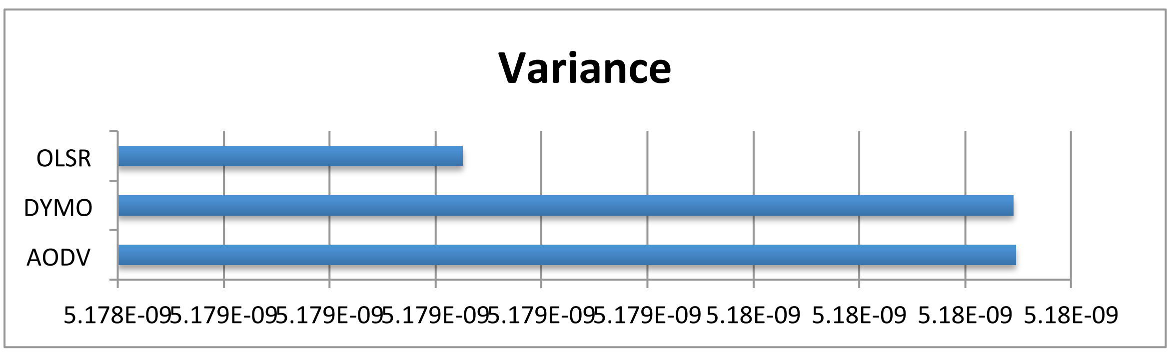

Figure 51.   Variance of Nodal Efficiencies by Protocol for LEO

The Hybrid constellation had similar efficiencies to the LEO, however, the efficiencies of the reactive protocols was boosted to the level of OSPF in the LEO while OLSR remained very low. This is captured in Figure 52.

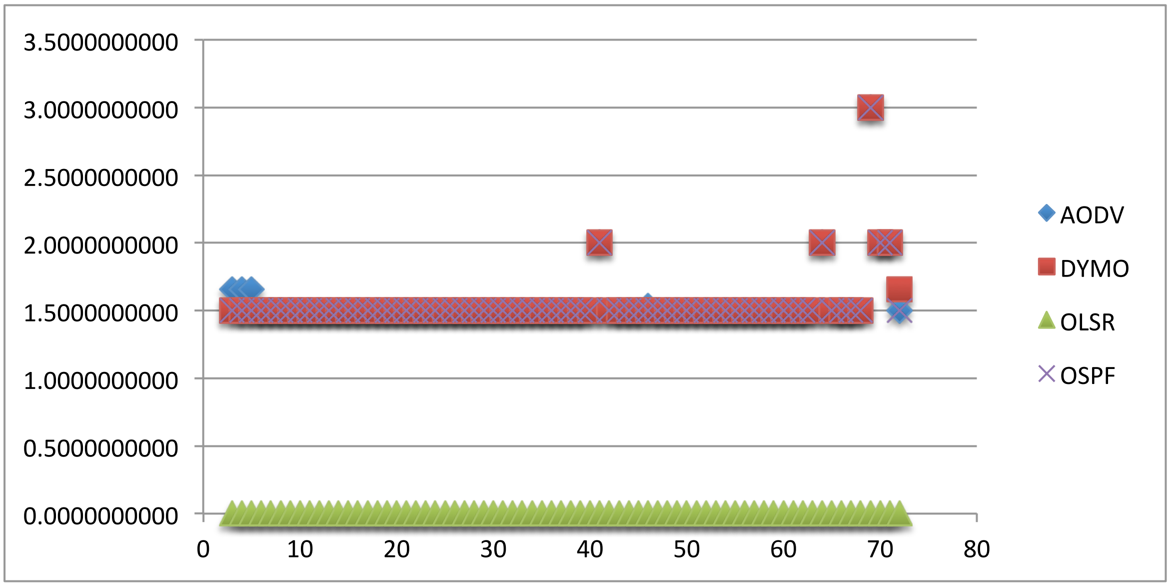

Figure 52. Node Efficiency by Protocol in the Hybrid Constellation

By zooming in on each group of efficiency data we can see that they are all about the same in the higher group with OSPF being the most variable. OLSR efficiency is much lower, but also very variable. This is captured in Figures 53 and 54.

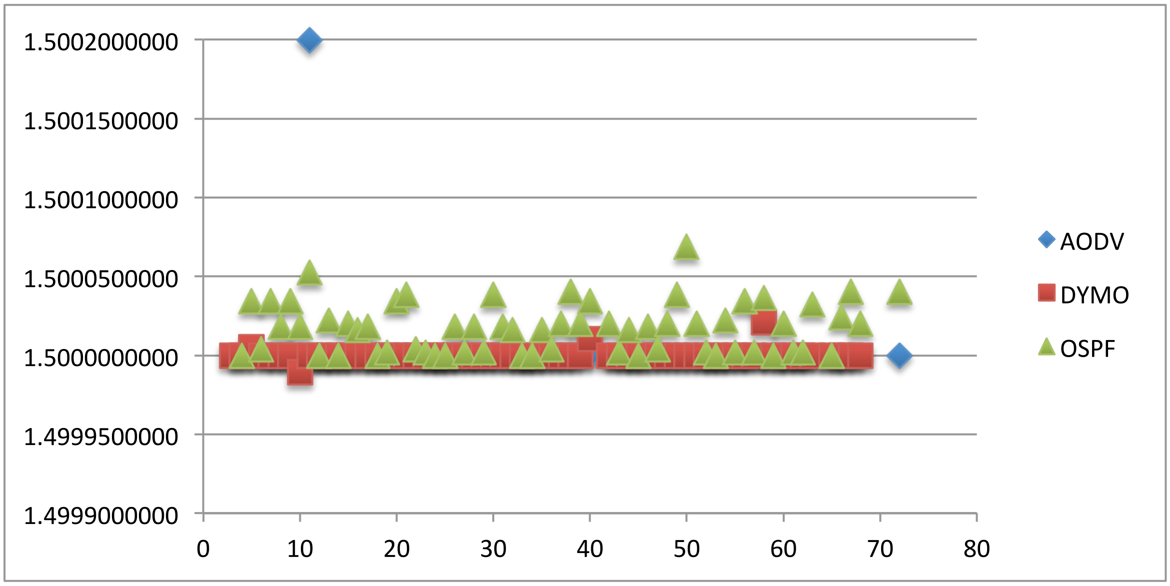

Figure 53. Node Efficiency by Protocol in the Hybrid Constellation v2

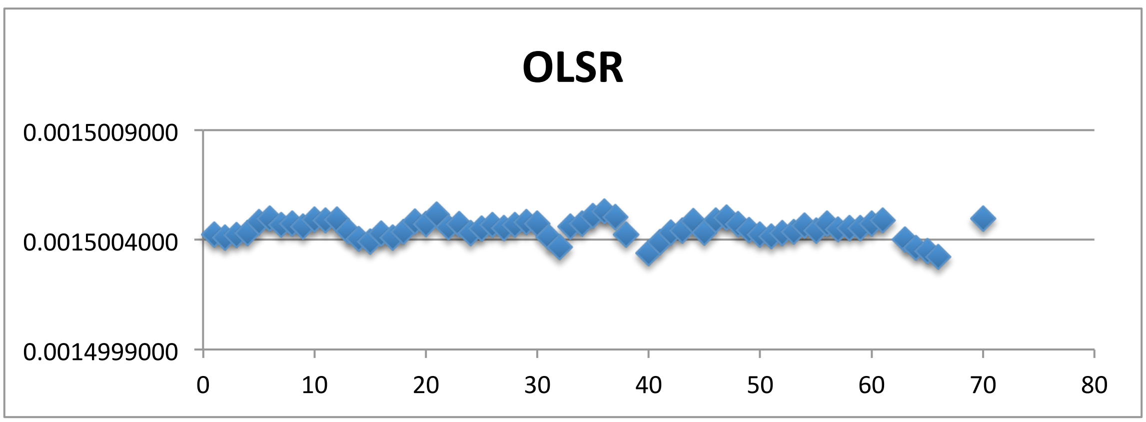

Figure 54. Node Efficiency of OLSR in the Hybrid Constellation

The variance graph capture in Figure 55 also shows the relative variation. However, the outliers in DYMO and AODV coupled with the relative scale of the numbers skews the data to make OLSR seem much less variable.

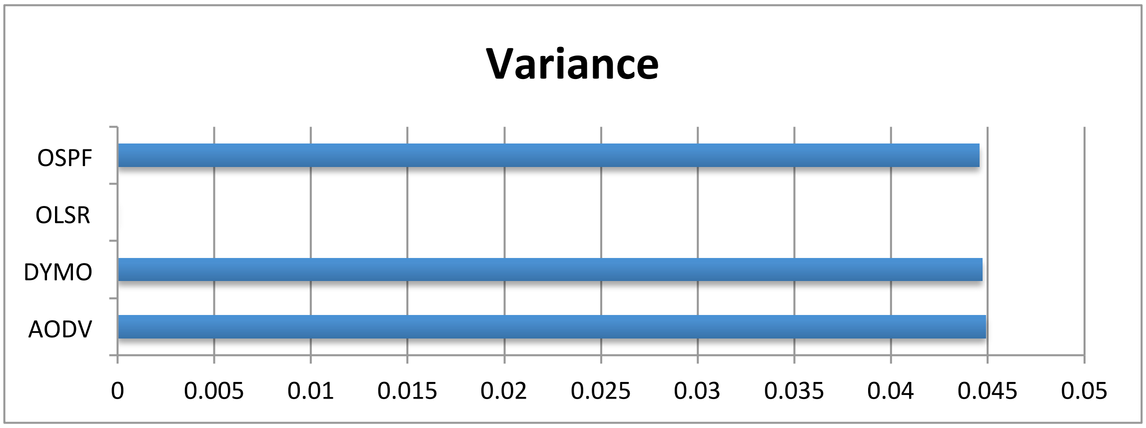

Figure 55. Variance in Nodal Efficiency of Protocols in the Hybrid Constellation

The Hosted constellation, being indicative of a more random architecture, has a wider variation in nodal efficiencies. The highest efficiencies are on the LEO satellites. This is due to the simplistic physical setup on each LEO satellite. Because they each only have a large beamwidth static antenna, they are rarely in contact with other satellites and so the ratio of IPin to IPout is high because this metric excludes any issues with signals being detected and locked on by the physical layer. As a result, the LEO satellites have nodal efficiencies around double that of the rest of the nodes. This is capture in Figure 56.

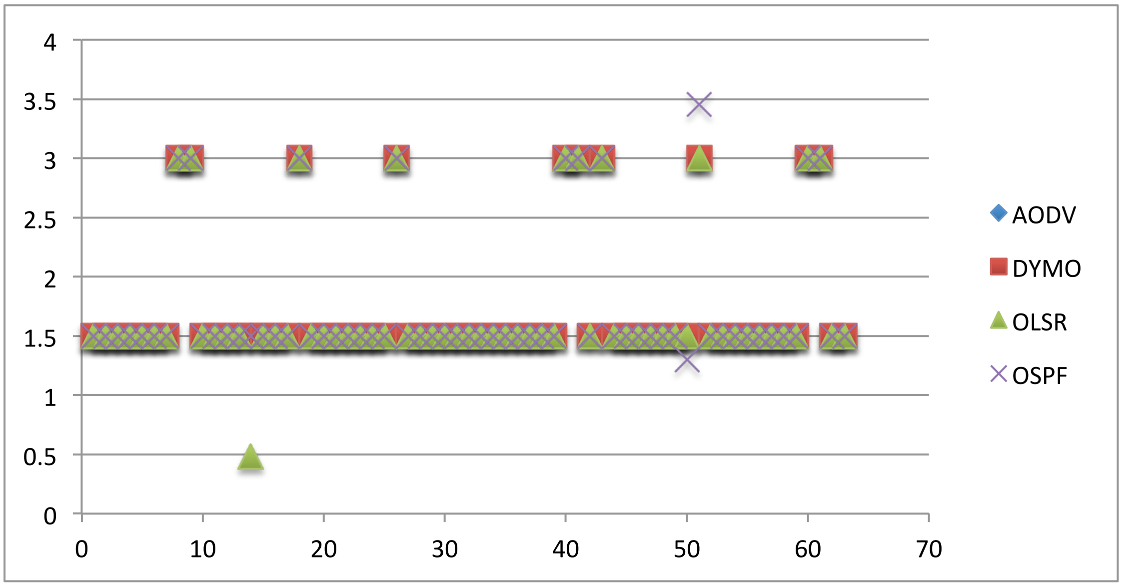

Figure 56. Node Efficiency by Protocol on the Hosted Constellation

By taking a closer look at the OLSR and OSPF data, it is clear in Figure 57 that the two had very similar efficiencies, but OLSR has much higher variation in data. AODV and DYMO are left out because they both have identical data and nearly no variation between nodes.

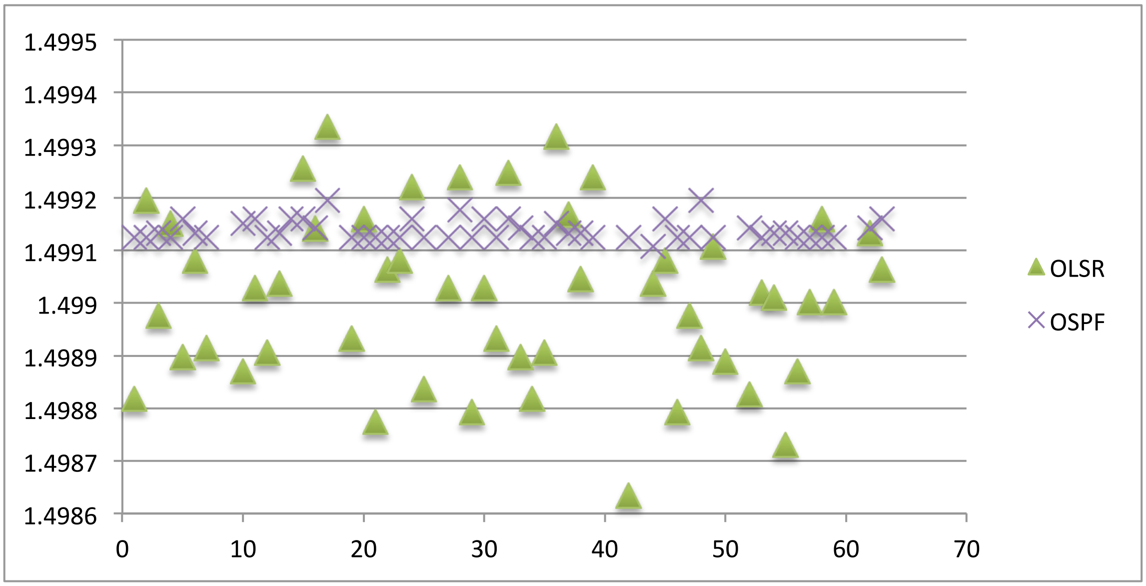

Figure 57. Nodal Efficiencies of OLSR and OSPF in the Hosted Constellation

Examining the coefficient of variation displayed in Figure 58, the standard deviation of the sample divided by the average value of the sample, for the different protocols in the Hosted constellation illuminates which protocol was more adept at efficiently dealing with the randomness of the Hosted constellation.

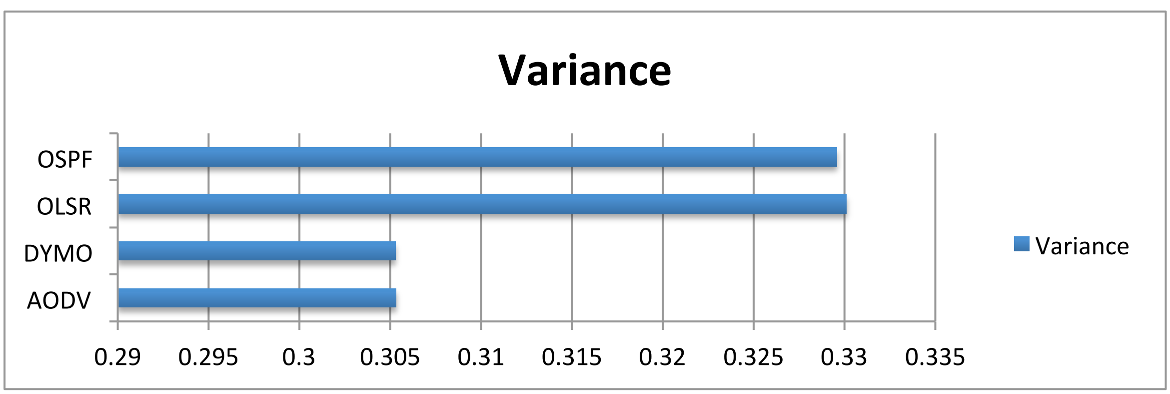

Figure 58. Variance by Protocol on the Hosted Constellation

2. Protocols by Constellation

The results of the modeling are also compared by constellation across each protocol. So for each protocol, the nodal efficiencies in each constellation are depicted in relation to one another, allowing relative conclusions to be drawn about the efficacy of the protocol across constellations as well as a comparative analysis of each constellation.

The AODV protocol has the best performance in the Hybrid and Hosted constellations with the lowest being in the LEO constellation. GEO is in between the two extremes. This is likely due to the GEO satellites in the Hybrid, Hosted and GEO constellations providing route data on their neighbor nodes in the LEO orbits or ground stations permitting the protocol to find routes quickly. In the LEO constellation with relatively few stable neighbors able to provide route data, more RREQ messages are generated, increasing network overhead and decreasing efficiency to a small fraction of those in a more structured constellation. This is captured in Figure 59.

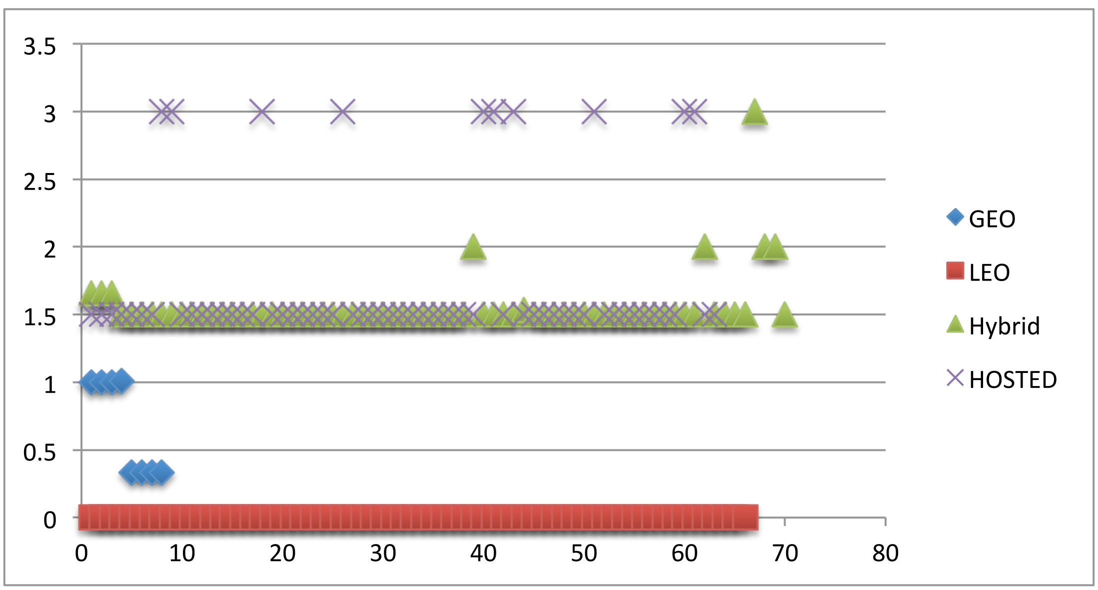

Figure 59. Nodal Efficiencies of the AODV protocol across Constellations

The average nodal efficiency by constellation clearly shows the relative suitability of AODV on each constellation, with Hosted being the most suitable. This is captured in Figure 60.

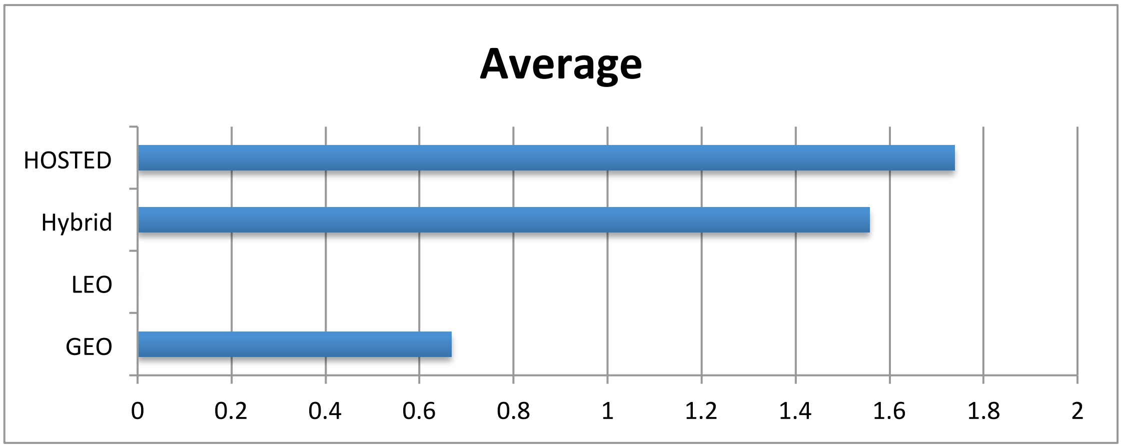

Figure 60. Average Nodal Efficiency of AODV by Constellation

The DYMO protocol, being very similar to the AODV protocol produced similar results to the AODV and for the same reasons. This is displayed in Figure 61.

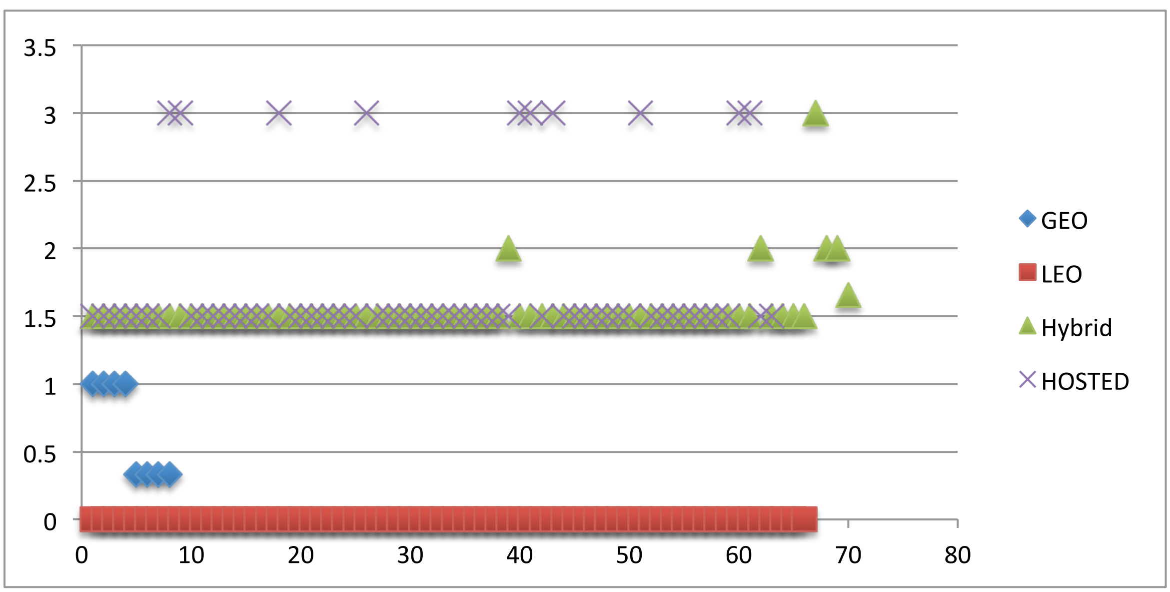

Figure 61. Nodal Efficiency of DYMO Protocol by Constellation

The average efficiency as displayed in Figure 62 shows the same suitability to each constellation as that of the AODV protocol. The performance of the on-demand reactive protocols increases as the number of relatively stable nodes increases. As a protocol for satellite networks, the reactive protocols are ill suited for dynamic orbital networks of homogenous relative stability.

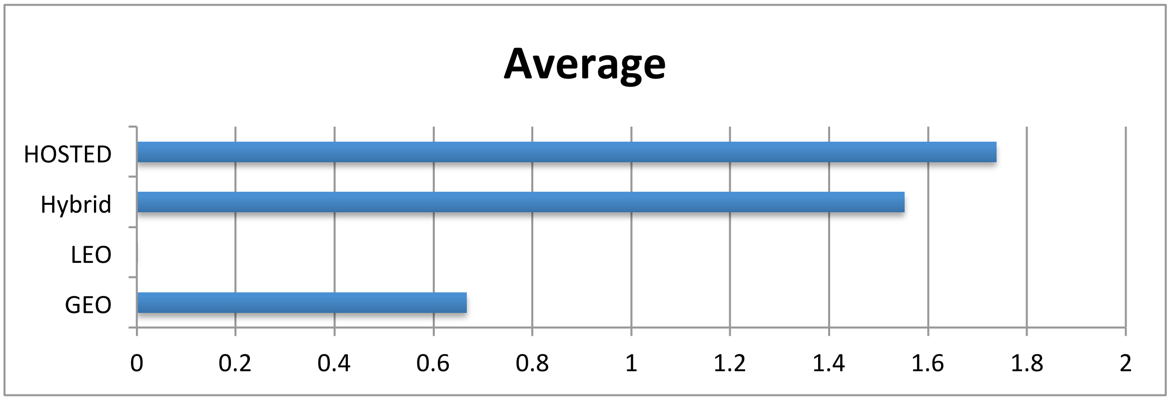

Figure 62. Average Nodal Efficiency of DYMO Protocol by Constellation

Of the two proactive protocols, OLSR displayed the greatest disparity among the various constellations, with the Hosted constellation performing significantly better than the other constellations. This is most likely due to the large number of GEO satellites in the Hosted constellation which provides a large pool of potential Multi Point Relays (MPR) and thus is able to attain a much higher nodal efficiency. This is captured in Figure 63.

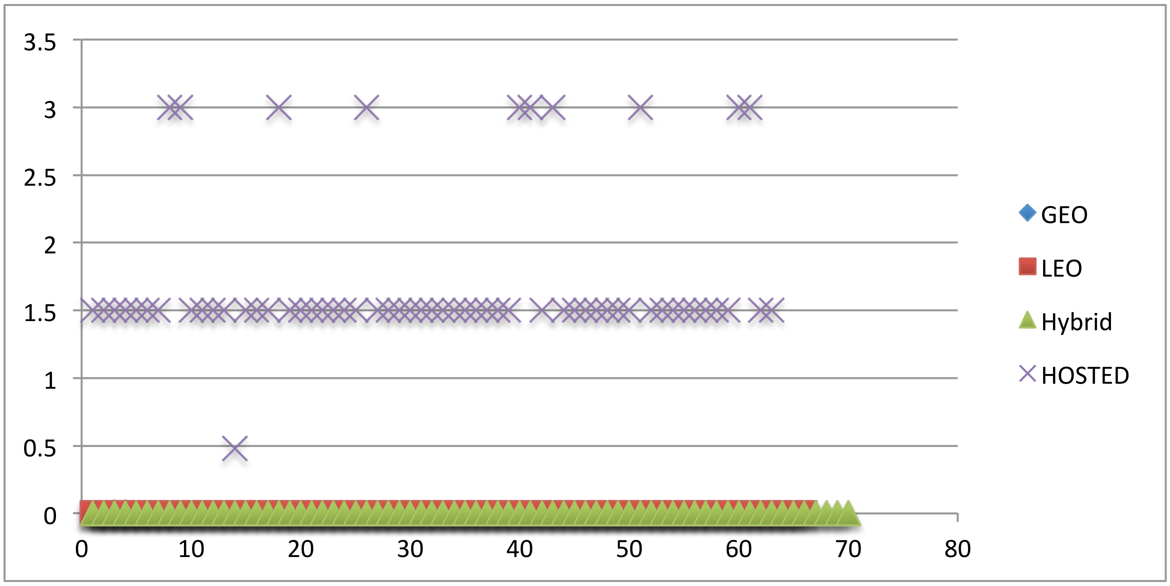

Figure 63. Nodal Efficiency of OLSR protocol by Constellation

A closer examination of the three lower efficiency constellations shows that the Hybrid Constellation performed better than the LEO and the GEO constellation is split, with the satellite nodes having better efficiency and the ground nodes having very low efficiency. This is captured in Figure 64.

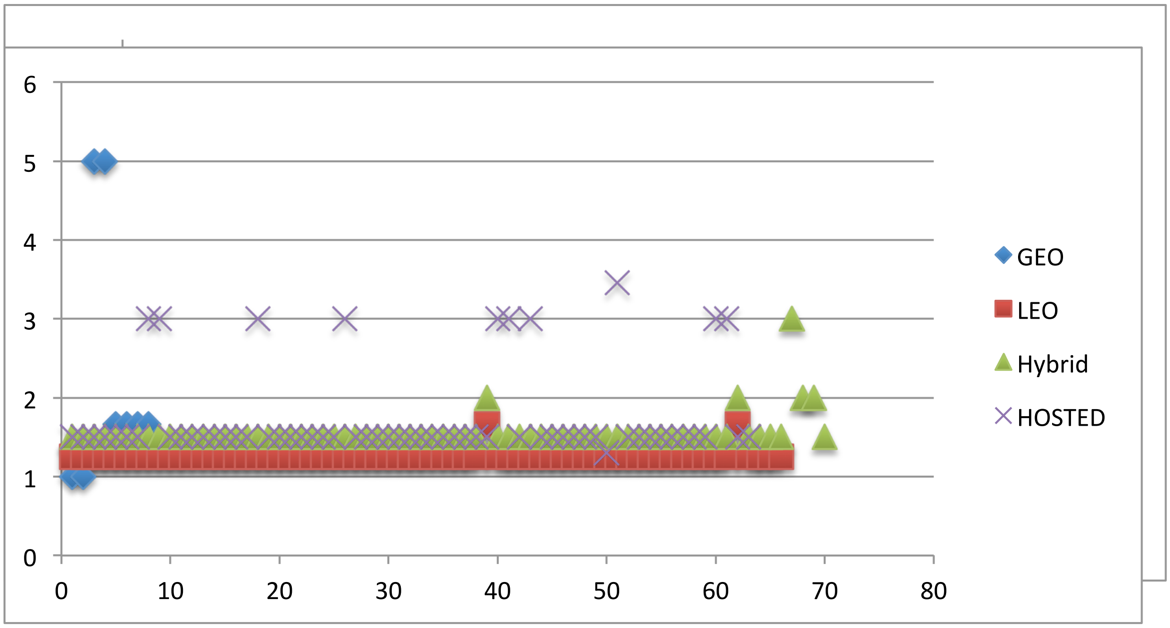

Figure 64. Nodal Efficiency of OLSR Protocol in GEO, LEO and Hybrid Constellations

The OSPFv2 protocol displays better average efficiency across the constellations and the efficiencies are all relatively close, indicating that OSPFv2 protocol works well regardless of orbital regime. The GEO transmitting nodes have the highest efficiencies of any of the data sets, and the rest are all high and consistent. Taking a closer look at the data shows that the Hybrid and Hosted constellations are very similar and that the LEO constellation is slightly less efficient. This is capture in Figure 65.

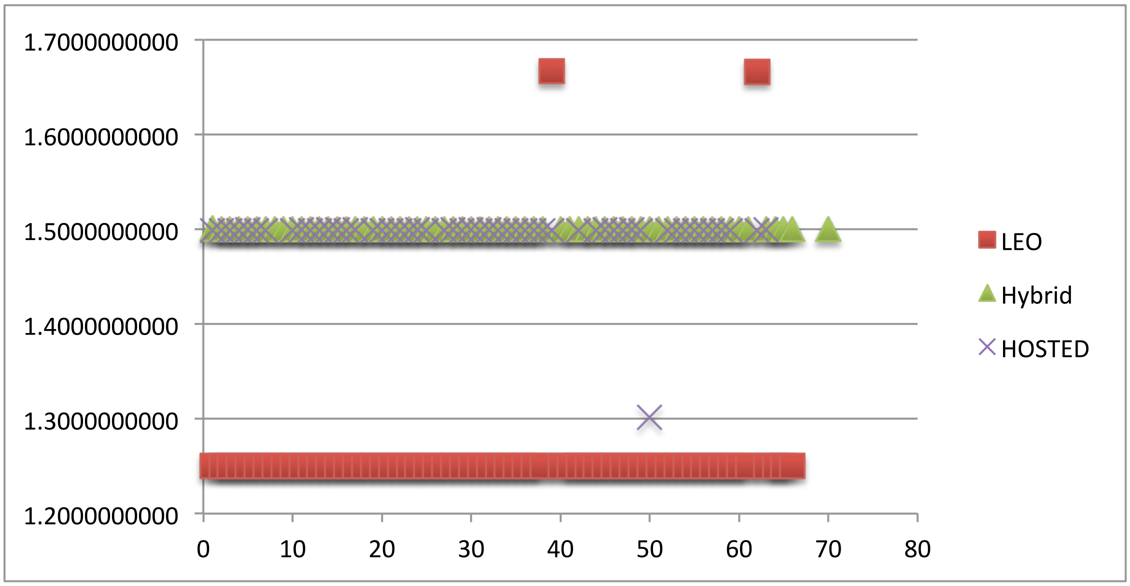

Figure 65. Nodal Efficiency of OSPFv2 in LEO, Hybrid and Hosted Constellations

The average efficiency of each node as displayed in Figure 66 shows that the GEO constellation has the best efficiency, followed by the Hosted constellation, the by the Hybrid and the LEO constellation is the least efficient.

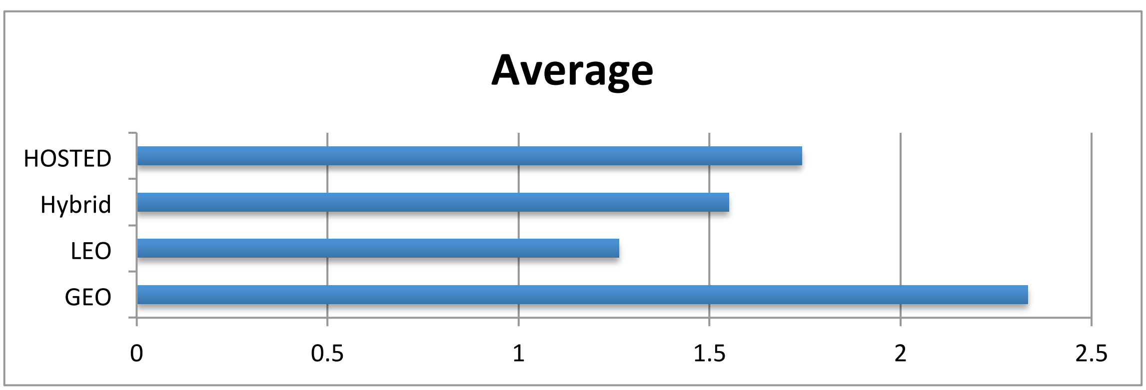

Figure 66. Average Node Efficiency of OSPFv2 by Constellation

C. THEORETICAL LATENCY

There are several factors that go into calculating the latency a signal experiences when being transmitted from point a to point b. The most obvious

ones are the total distance traveled, speed at which the signal travels, processing time experienced at each node as well as time the message spent in the queue at each node. The latency experience for a signal traversing a GEO constellation is longer than a signal traversing LEO constellation. The general formula for calculating latency is given in the equation below. Where T_t is the total time it takes for the signal to be transmitted from the transmitting station to the receiving station, D_t is the total distance the signal traverses, c is the speed of light, n is the total number of satellite nodes the signal traverses, T_{queue} is the time the signal spends in the queue and $T_{processing}$ is the processing delay of the node. The processing time will be the same for every node regardless of the routing protocol used, while the queue time will vary from one routing protocol to another.

$$T_t = \frac{D_t}{c} + n\left(T_{queue} + T_{processing}\right) \quad (0.19)$$

The process for determining the total distance traveled is slightly different between the LEO and the GEO constellations, therefore that variable will be defined in greater detail in the following subsections.

The method used for estimating the queue time was pulled from data obtained when the protocols were modeled using Qualnet. One of the statistics Qualnet provides is the longest time in queue. This is the statistic that was used estimate the time a packet spent in the queue in a particular node. To estimate the processing time experienced by a node, the processing time was estimated to be 25% of the longest queue time. Because the processing time could not possibly be longer the queue time, it was determined the best approximation for the processing time was 25% of the longest queue time. This same process was used for each protocol in both the GEO and the LEO constellations.

1. GEO Latency

The approach taken to answer this question was to derive the longest distance the signal would need to traverse. In the case of the GEO constellation, where the satellites are equally spaced out at 90 degree intervals along the equator, the longest any signal would need to travel would be in cases where the difference between the transmitter’s and the receiver’s longitude is 180 degrees; meaning they are on opposite sides of the earth. It is assumed that both the transmitter and receiver are directly beneath GEO satellites along the equator. This means the angle between two satellites measured at the center of the earth is 90 degrees, bisecting that angle creates two 45 degree angles resulting in the final angle also being 45 degrees. Drawing a line that bisects the earth central angle and is perpendicular to the transmission path between the two satellites gives a distance of 42,056 km, forming an isosceles triangle. Note this is equal to the sum of the radius of the earth and the altitude of the satellites. Because it is an isosceles triangle the distance between two GEO satellites spaced and 90 degrees longitude along the equator is twice this distance or 84,112 km. Figure 67 describes how this distance was calculated. This method was confirmed by using Satellite Took Kit to model the constellation. Next STK was used to display the fixed position coordinates x, y, z and the distance equation was used to verify the distance. Figure 68 is a 2D representation of the GEO constellation. Figure 69 captures the x, y, z coordinates of the satellites. Table 9 contains estimated latency of a signal traversing the GEO constellation.

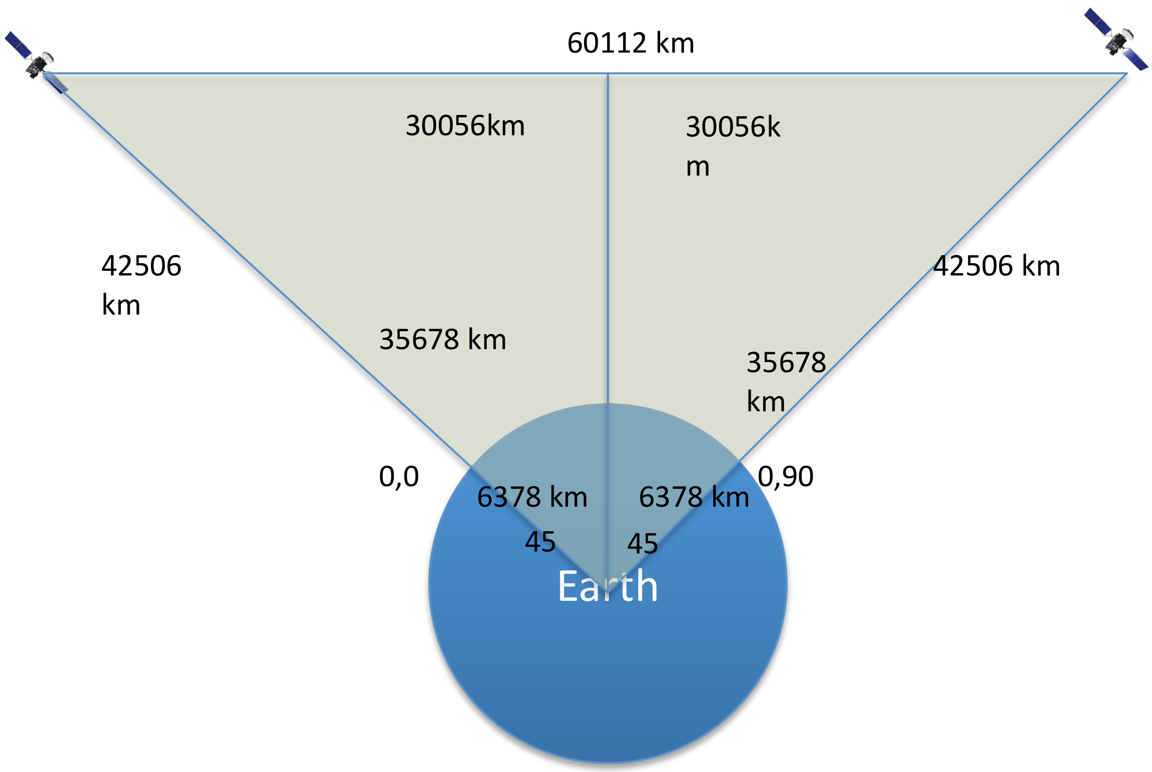

Figure 67. Distance Between GEO Satellites

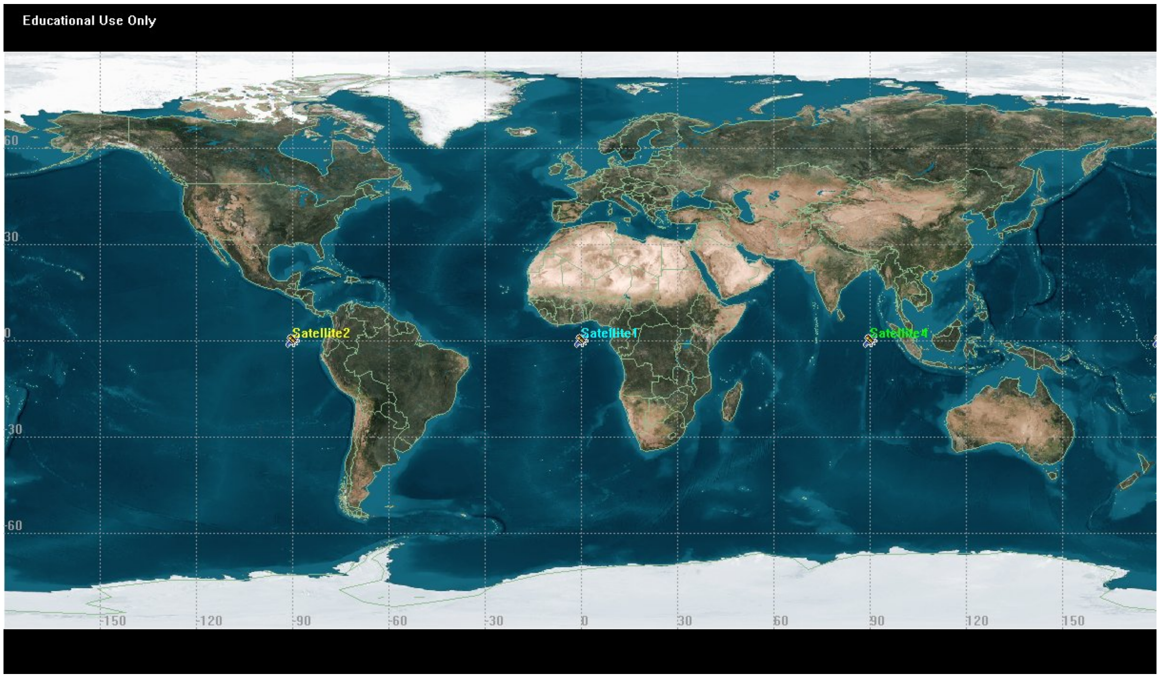

Figure 68. GEO Constellation in 2D

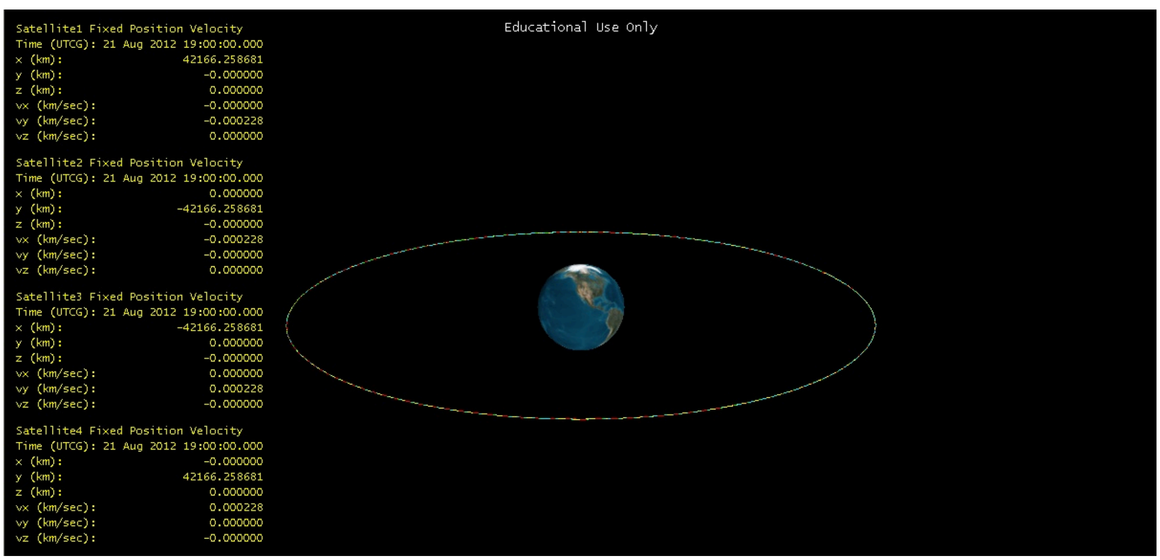

Figure 69. GEO Constellation with Fixed XYZ

	Protocol			
	OLSR	AODV	DYMO	OSPF V2
Distance (km)	59632.09394	59632.09394	59632.09394	59632.09394
Total Sats	3	3	3	3
Queue (s)	0.012	0.012	0.012	0.012
Processing(s)	0.005	0.005	0.005	0.005
Latency	0.249911255	0.249911255	0.249911255	0.249911255

Table 9. GEO Theoretical Latency

Now that the distance between two GEO satellites spaced out at 90 degrees along the equator has been calculated, D_t can be more clearly defined for this instance. Because the transmitter and receiver are 180 degrees out from one another, a total of three satellites must be used, therefore, n=3. It is also known that the signal must travel from the surface of the earth to a satellite (transmitter) and then from a satellite to the surface of the earth (receiver), in addition to distance it must travel between satellites. The equation below describes how to calculate the total distance the signal must travel, where H is the altitude of the satellite and $D_{satellite}$ is the distance between the satellites.

$$D_t = 2H + (n-1)D_{satellite} \qquad (0.20)$$

Substituting equation 0.20 into equation 0.19 yields:

$$T_t = \frac{2H + (n-1)D_{satellite}}{c} + n\left(T_{queue} + T_{processing}\right) \qquad (0.21)$$

2. LEO Latency

While the general formula for calculating the general latency remains fairly close to that for the GEO constellation, there are some differences that are worth noting. For example, while the GEO satellites are spaced out at 90 degree intervals along the equator the same is not the case for the LEO constellation. While the LEO orbital planes are spaced out evenly across the equator, the

satellites themselves are also phased evenly to ensure there is constant coverage within the plane. Figure 70 is a screenshot taken from STK that displays the sub satellite points on the surface of the Earth. This was particularly useful to determine a notional path between two stations on the surface of the earth whose difference between longitudes was at approximately 180 degrees. Because the satellites are phased within their planes, satellites in adjacent planes are not directly across from one another. A notional path was determined beginning at Iridium 76, continuing through Iridium 29, 35, 52, 40, 86 and finally ending with Iridium 20. In order to calculate the distance between the satellites, STK was used to display the x, y, z fixed position coordinates, this is captured in Figure 71. Once the coordinates were obtained, the distance equation was used to calculate the distance between the satellites. The general form of the distance equation is expressed in equation 0.22 Table 10 contains the latency experienced by the protocols examined in the experiment.

$$D = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2 + (z_1 - z_2)} \qquad (0.22)$$

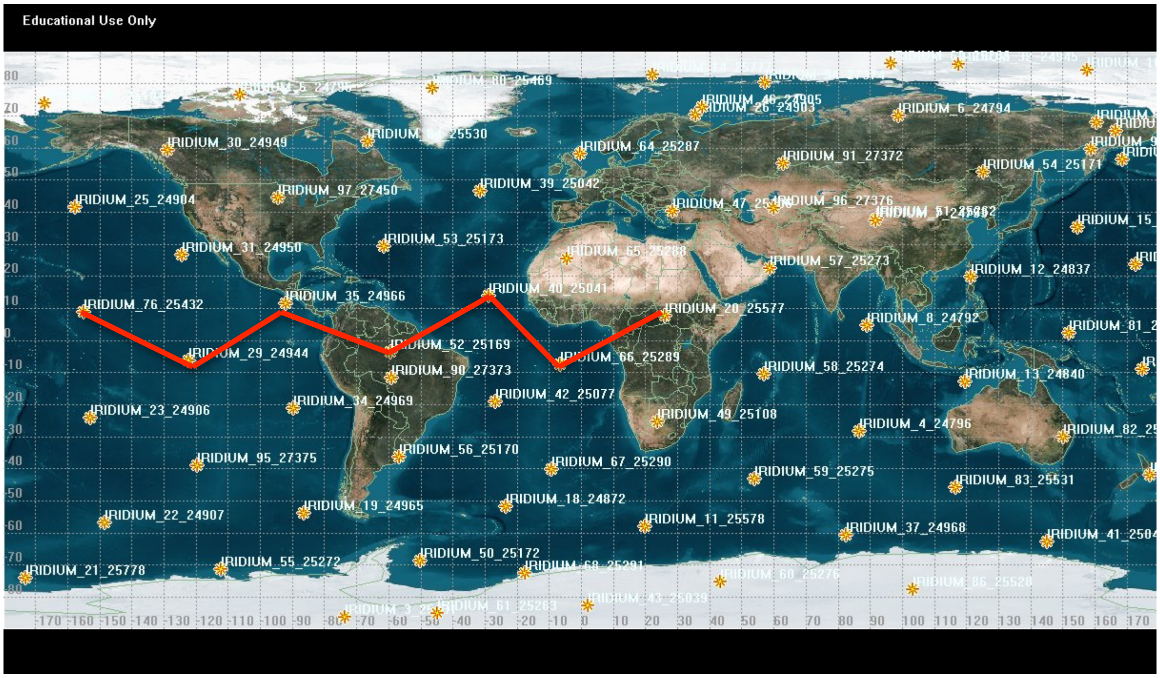

Figure 70. Iridium Notional Path

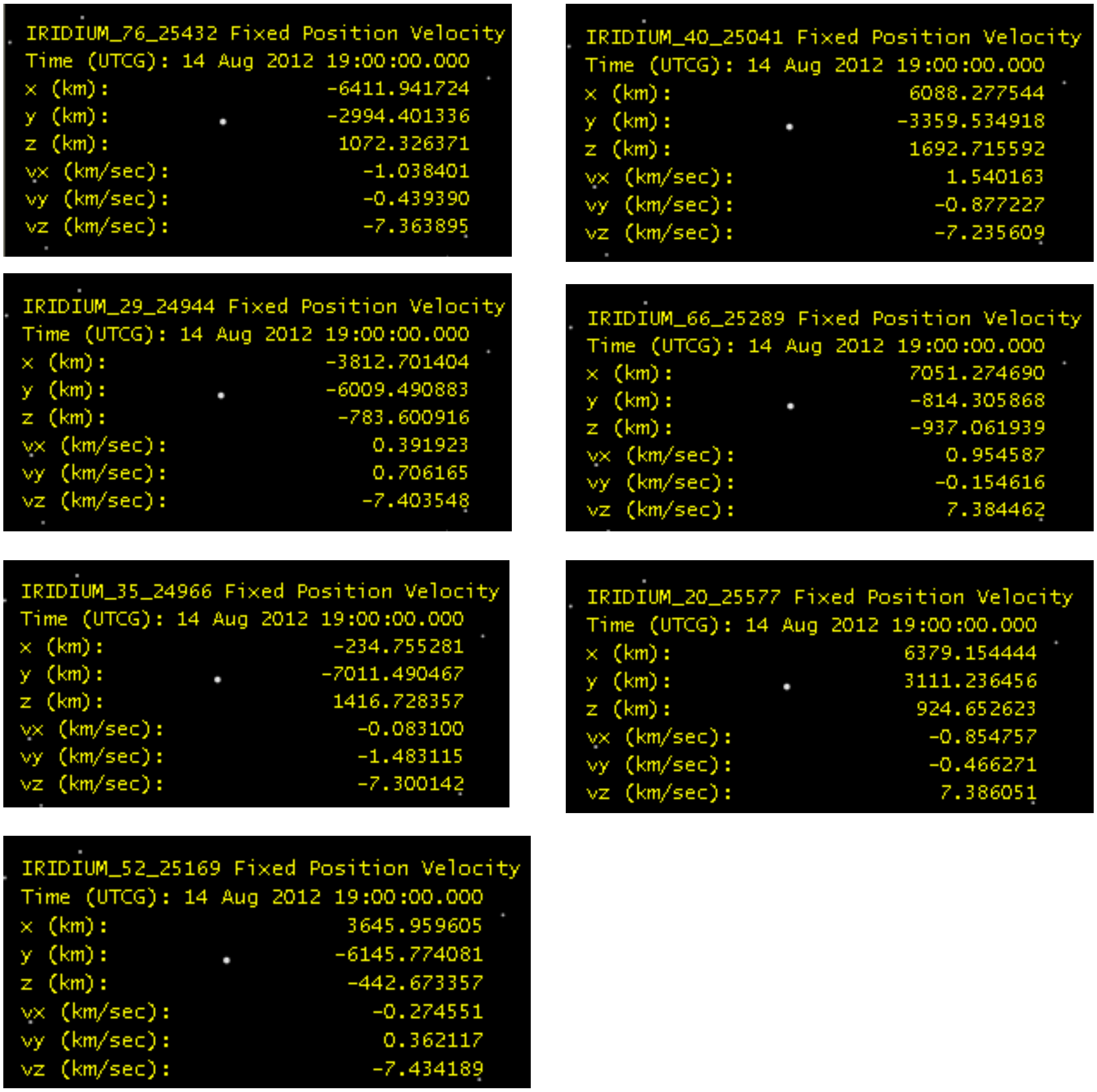

Figure 71. x, y, z Coordinates From STK

	Protocol			
	OLSR	AODV	DYMO	OSPF V2
Distance (km)	27096.9179	27096.9179	27096.9179	27096.9179
Total Sats	7	7	7	7
Queue (s)	0.012	0.012	0.012	0.012
Processing(s)	0.005	0.005	0.005	0.005
Latency	0.209385589	0.209385589	0.209385589	0.209385589

Table 10. LEO Theoretical Latency

The theoretical latencies do not substantially change between constellations and do not change at all between protocols. In terms of relative merit, the LEO constellation has a slightly lower latency than the GEO constellation and the Hybrid and Hosted constellations would be in between. This information is useful in formulating the appropriate delay tolerance for a space specific network. The latency does not appear to be affected by protocol.

D. AVERAGE HOP COUNT

The hop count, as calculated by QualNet, by protocol and by constellation indicates that this number is generated by some mathematical process within QualNet that is not conclusive for comparison of orbital networks, either constellations or protocols. The OSPFv2 hop count was identical in every constellation, as unlikely as that would seem while the GEO constellation seemed to have the highest number of hops despite the fact that the GEO constellation has the fewest nodes and therefore the lowest possible number of hops. The data is presented in Figure 72 for completeness, but is ultimately inconclusive.

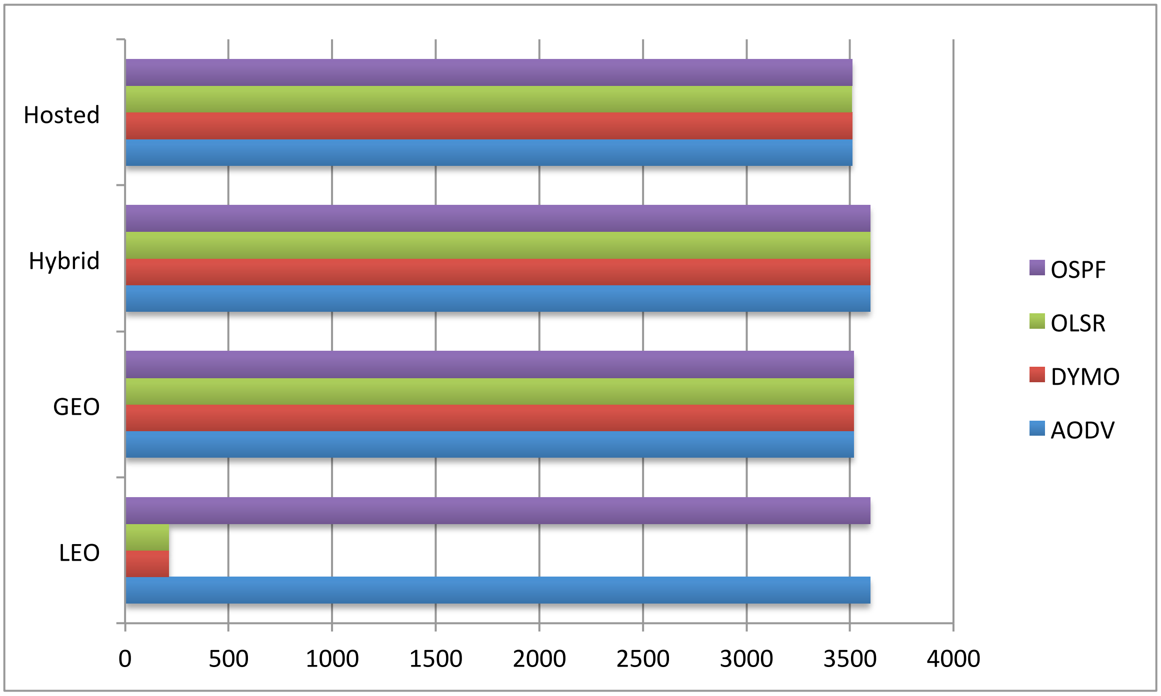

Figure 72. Average Hop Count by Protocol and Constellation

E. PACKETS DROPPED FOR NO ROUTE

The number of packets dropped for no route in each protocol and each constellation as displayed in Figure 73 shows very little variation. This is in part due to the fact that the packets dropped is highly dependent on the number of data packets sent, and in this scenario, every simulation consisted of a single node sending a finite number of packets. It is also possible that this could be due to effects occurring at the Physical Layer that would appear constant across each protocol simulation for a constellation. The numbers are fairly constant by constellation and ultimately provide no conclusive information. Of note, is the difference between the protocols with respect to the LEO constellation in which OLSR and DYMO had significantly less packets dropped, but could not be attributed to a property of the protocol.

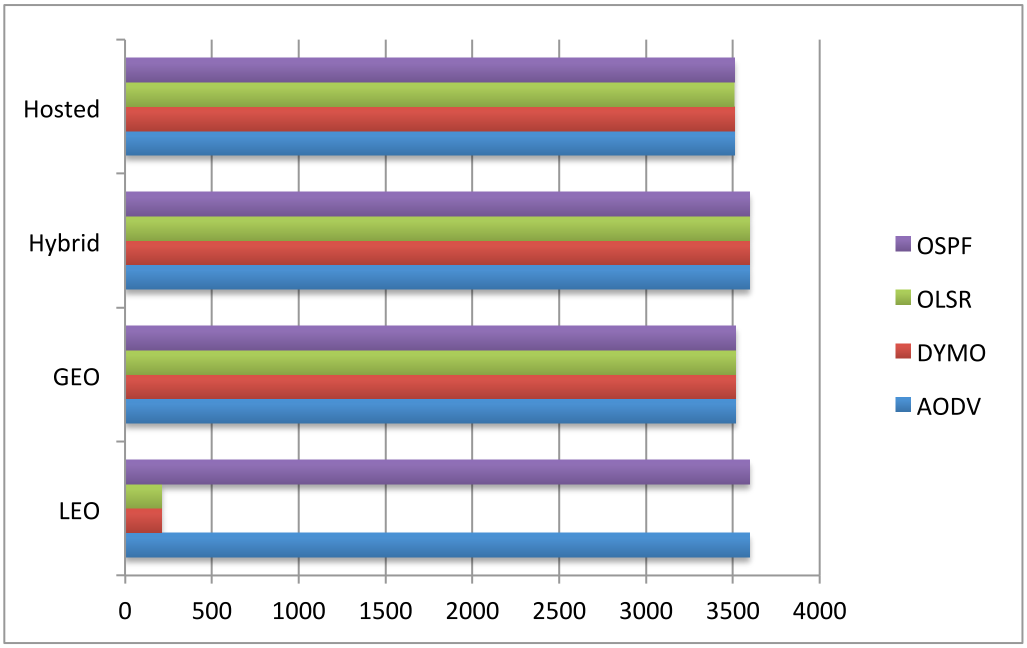

Figure 73. Average Hop Count

F. CONCLUSIONS

Each protocol and each constellation, being analyzed in terms of efficient use of network bandwidth, returned results that would tend to suggest that OSPFv2 is marginally better suited to the space environment's inherent mix of static and dynamic links than other protocols tested. OSPFv2 operates by flooding the network with discovery messages before generating a single network topology that all routers agree on and generate routes based on. Given the constraints on simulation time that this study was conducted under, the models of each network were run for a short enough timeframe that a single network discovery was sufficient to make routing decisions. However, given longer time scales of simulation, it is very likely that the efficiency of OSPFv2 would converge on some efficiency that is much lower than those represented by the data in this study.

In all other measures of effectiveness or suitability, there is no clear "best choice" for a space based network protocol. If any of the tested protocols were

well suited to the space network, then the data would show conclusively that that protocol was best in several areas. However, given the number of inconclusive results and in some cases, inconsistent results, it is evident that none of the protocols are well suited to the space based network.

A network of satellites operates in a very unique combination of predictable static links and variable dynamic links in such a way that protocols designed to operate in static networks have trouble managing the dynamic links and protocols which are designed for dynamic networks, such as MANET's, are overly inefficient for the level of dynamic links in the network. Satellite networks also have the unique attribute of being predictably dynamic because their motion is based on orbital mechanics and are thus extremely predictable within their specific orbital regime. The dynamic links arise when links are made between satellites of differing orbital regimes. A protocol optimized for satellite nodes should be able to account for this behavior by predicting changes in topology due to orbits and adapt accordingly.

THIS PAGE INTENTIONALLY LEFT BLANK

VII. SPACE ROUTING PROTOCOL DESIGN

A. SOCIAL NETWORKING

1. INTRODUCTION

In order to fully utilize the network centric type of communications envisioned in this thesis, established routing protocols were used in comparing the different satellite constellations. These protocols include proactive and reactive protocol types that are currently utilized in mobile ad-hoc networks as well as a baseline protocol used on many wired networks. The field of social networking has introduced a new paradigm into the analysis of networking theory. Social networking theory seeks to explain how the relationships or ties that exist between individuals within a group affect the flow of information and change the way both the group and the individuals act. (Wasserman 1994) By applying some of these theories to networking protocol design, a new protocol model is realized that can be adapted for the unique space networking environment.

2. Adaptation

a. Definition

The concept of adaptation displayed in Figure 74 appears to be a fairly simple concept. If you were to ask a person at random to define adaptation, you are likely to receive something close to the correct answer. Holland defines it by saying, "[Adaptation] is a study of how systems can generate procedures enabling them to adjust efficiently to their environments." (Holland 1962). This definition involves the interaction between an entity and its environment and it includes qualitative assumptions about the process by using such words as "efficiently." And while this definition is sufficient for *defining* and *identifying* adaptation, it is wholly insufficient for the purposes of *designing* adaptation into a complex system.

Adaptation is present in most emergent complex systems, but setting out to create a complex system that automatically and intelligently adapts has proven extremely difficult. In order to set upon the task of designing adaptation into a system, the process of adaptation must be examined and defined in such a way that it may be reproduced. We will show that the process of adaption is comprised of six steps, each step being a prerequisite to the next and that each step adds additional levels of complexity to the system.

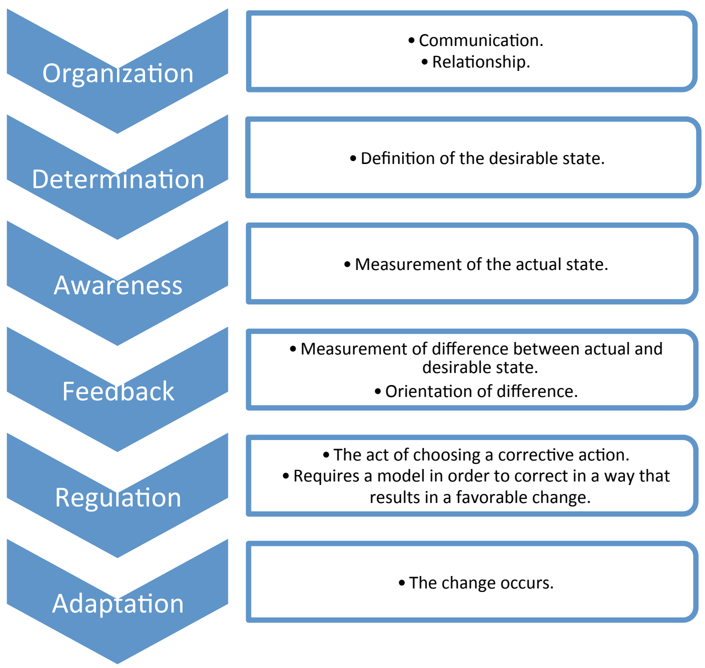

Figure 74. Adaptation Process

The first requirement for adaptation is the presence of organization. The best definition of organization is provided by W. Ross Ashby as a conditionality between interrelated entities, and that as soon as a relation between two entities exists such that one's state is conditional on the other's state, a component of organization is present (Ashby 1962). In this understanding of organizations, it is important to note, as Ashby points out, that the presence of organization makes no claims to the relative value of the organization. Whether the organization is a "good" one or a "bad" one is

immaterial to the definition of organization for our purposes. If two entities exist in perfect apathy to each other such that no relations can be drawn between them, or if two entities exist in perfect isolation such that no medium exists for them to communicate then there is no organization. Therefore, organization must exist and can only exist between any two entities that are both capable of communication and have a relationship between them. If two computers are designed to operate in concert but have no medium for communication, then no organization can exist since a change of state in the first computer can cause no change in the second computer.

The nature of this relationship must take the form of a natural law so that when a change in A is communicated to B, a resulting change in B must always take place and the change must be such that an infinite number of perturbations under identical circumstances will always result in the same change in B for any one change in A. This can be as simple as an array of hard drives that are set up to mirror data, or as complex as the interaction between the flux density of cosmic rays and the Earth's cloud cover. The point is that the interaction follows natural laws that allow results to be reproduced, or predicted.

b. Determination

The second step in the process of adaptation is determination. Ashby set out in 1962 to describe the principles of self-organization. He built upon the concept of organizations and defined two ways in which a system can be "self-organized." The first method of self-organization is the process of communications and relations forming between entities. As many individual and separate parts form relationships between each other, an organization emerges. This is the process that we understand today is responsible for the coalescing of life out of constituent chemicals. It is random, or at least, not directed and the only qualitative measure of such systems is whether they continue to exist. In natural systems of organisms, what is good or bad is simply the difference between what is replicating and what is dead. What biological systems and

artificial systems that adapt share is that there is a definition of success. In an organism, the definition of success is continuing life. In an artificial system, success is a set of design constraints. Regardless of how success is defined and whether it is a natural emergent definition or a programmed objective, the act of defining the goal of the system is a critical step in the process of adaptation.

c. Feedback

The third and fourth steps are closely intertwined and together constitute traditional feedback. Step three is awareness. It is the ability to measure the actual state of the system. The fourth step is feedback, which Ramaprasad defines as "information about the gap between the actual level and the reference level of a system parameter." (Ramaprasad 1983). Feedback represents the measurement of the orientation and difference between the desired state and the actual state. As an easy example, assume that you must monitor the temperature of water in a bathtub. Your definition of the desired state is "warm" and a thermometer placed in the tub measures your awareness. In order to consider your awareness as feedback, you must be able to translate the measured temperature into some state relative to your desired goal. If you are unable to determine if 80 degrees Fahrenheit is above or below "warm," then you will be unable to classify such as feedback. Conversely, if your desired state is 80 degrees F, and your thermometer reads 75 degrees F, then you know that you must increase temperature. Therefore, for our purposes, feedback is the process by which measurement of actual state is translated into a form that is a relation to the desired state. If no relation can be drawn between desired and actual state, then no feedback exists.

d. Regulation

The fifth step is regulation. Conant and Ashby, through rigorous logical proof concluded that in order for a regulator to succeed in regulating a system, the regulator must be a model of the system (Ashby and Conant 1970). The basic premise is that in order for a regulator to take action in the process of

regulation, one must have a good idea of what actions will result in what outcomes. Therefore, whether preprogrammed into a system or built into its design, or injected with human interaction, a regulation to return a system to its desired state is predicated on the ability to model the outcomes of various actions. We have defined regulation as the act of choosing corrective action in response to feedback and so doing, requires a model to understand *how* to correct in a way that results in a favorable change.

As example, take the same bathtub. If your feedback is that the temperature must increase by 5 degrees, you would instinctively reach for the hot water tap, but assume for the sake of the example, that the taps are not labeled and so you do not know which is hot and which cold. You would have to test each to find out which one is hot. To add a level of complexity, imagine that you must do so remotely. So now you must run one faucet for a few minutes and observe the change in the temperature of the bathtub. You can measure the change, and then deduce which faucet is hot and which cold. This method is simple, and yet you should be able to get the temperature close to where you want it to be. If you were able to calculate such things as the heat transfer from the water to the air, the flow rate of the pipes, the heat transfer from the pipes into the walls, the amount of energy being input into the water in the hot water heater, and any number of other factors and could calculate the temperature of the tub for either faucet running for any time, then you would have a much better regulator of tub temperature.

The difference between the quality of regulation is the fidelity of the model. The first example, just knowing which faucet was hot and cold represents a model that tells you the direction of a change vector in the temperature of the tub, but not the magnitude. It is also a binary control, on or off. The second is a much more complex model and truer to real life and thus makes a better regulator of the bathtub temperature system.

The simplest regulator is a binary control with feedback system and only knows one action to correct for deviation from the desired state. A more

complex regulator might have a dozen or more available actions available to it and might be reacting to multiple forms of feedback but is still reacting. The most complex regulator is one that predicts changes in the environment and then models various actions to see which one will result in the appropriate action. This is how the human brain operates.

e. Adaptation

The sixth and final step is adaptation and it is the act of choosing a state and making the change. The act of adaptation can take one of three modes. As defined by Chakravathy (Chakravarthy 1982), the three modes of adaptation are defensive, reactive and proactive. Each of these modes requires a different level of complexity in the overall system and in the decision-making process, whether human in the loop or automated regulator. The modes of adaptation are captured in Figure 75.

The most basic level of adaptation is defensive. This mode of adaptation attempts to insulate the system from the effects of the environment. Much of the actual regulation is done during the design of the system and emphasis is given to reliability in any anticipated environment. This is the kind of adaptation that is predominant in systems engineering and military systems. Much thought is given to the types of environment that the system might encounter and then the system is designed to avoid being affected in any detrimental way by that environment. This is also how modern satellite communications are designed. In a dynamic system such as the corporations that were the focus of Chakravathy, this mode of adaptation corresponds to a defend and hold mentality in business. If a business is not concerned with enlarging their markets or developing new products, then they are likely to focus on keeping what they have.

Reactive adaptation is more complex, and is the predominant mode of adaptation in automated regulatory systems. When a change in the system is detected, a response is automatically generated to bring the system back into

acceptable parameters. These systems do not attempt to predict changes, they merely react to changes. The strength of these systems is that they are relatively simple and are predicable. The weakness of these systems is that they are unable to respond unforeseen changes in the environment. The bathtub example form earlier would be completely unable to respond to a loss of the water heater because such a thing is outside of its parameters.

Proactive adaption is the most complex. This mode of adaption requires models of high fidelity and the ability to model various actions to determine the best course even in situations that may not have been foreseen during design. These systems react to changes in the environment before they happen so that no loss of effectiveness of the organization is felt.

These principles of adaptation are appropriate whether discussing a communications system or a business. In terms of complexity, the systems that include humans are inherently much closer to operations like these simply because the human brain operates in this manner. When attempting to develop an automated system capable of intelligent adaptation, there is a much higher level of complexity involved. These kinds of systems will require a *feed-forward* system of proactive adaptation. This requires the ability to: sense or predict changes in the environment *before* they affect the organization, predict the effects of those changes, model changes in the organization to determine the outcomes of various courses of action, and the abilty to select the “best” adaptation to the situation.

f. Application

Based on Chakravathy’s work in defining modes of adaptation, the following model is proposed as a basis for routing decision making in satellite networks.

Nature of Relation to Environment	Organizational Fit	Routing Decision Scheme	OSI Layer
~~Defensive~~	Mechanistic	Transponded/ Bent Pipe	Physical
Reactive	Bureaucratic	Radio Aware (RAR)	Physical
Proactive (Predictive)	Organic	Orbit Defined	Physical
Static	Hierarchal	User Defined Priority Based	Application

Figure 75. Modes of Adaptation

In this model, the Defensive mode of adaptation is dismissed as the traditional transponded satellite network model in which routing decisions are made on the ground. Instead a new, 3 stage, routing decision model is proposed based on Reactive, Proactive and Static modes.

The first stage is the Proactive or predictive mode. This stage of the model would update the routing table based on the expected position of the satellite using a sliding frame finite state machine. Once a satellite constellation is in space, the orbits are well defined and the crosslinks between the nodes can be predicted for any given time period.

The second stage is the Reactive mode. Based upon an expected state for each link from stage one, a metric would be used to identify which links are behaving as expected and adjust the routing table accordingly. It would also allow adjustments to be made to transmitter power or bandwidth as require to maintain link quality.

The final stage would be based on application priority as defined by the user. In this mode, routing decisions would match high priority packets with higher quality links and vice versa.

3. Strong and Weak Ties

a. Foundation

The definitive work on social networks and the interaction between the micro and macro levels is Mark Granovetter's *The Strength of Weak Ties* (Granovetter 1973). He defines Strong Ties as the connections within small groups of people. In other words, if subject A is in constant or close contact with subject B then there exists a strong tie between Subject A and B. He classifies these as people you have almost daily contact with i.e., family, close friends and co-workers. He goes on to define weak ties as connections between groups of people. For example, as displayed in Figure 76, subject A is a member of group A and subject C is a member of group C. There does not exist a direct tie between subjects A and C. There does however, exist a tie between subject B (a member of group A) and subject C. Therefore, there is a weak tie between groups A and C. It is these weak ties between groups that are the strength of the network and facilitate the flow of information from group to group.

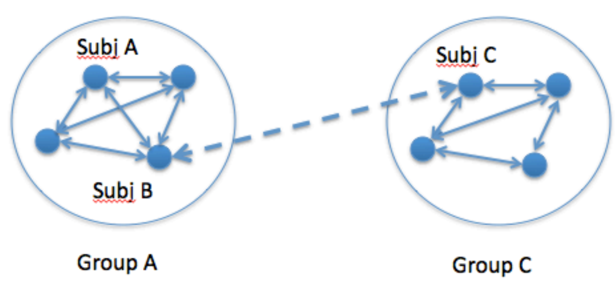

Figure 76. Strong and Weak Ties

Another example of the strength of weak ties was a study conducted by Stanley Milgram in 1967 in which he set out to define the social

distance between two randomly selected people in the United States. His experiment was simple in design. First, two people were selected at random, one a resident of Wichita and the other a resident of Omaha. The directions were simple. The person in Wichita was given a folder with some basic information about the target in Omaha and they were instructed to mail the folder to either the person in Omaha, or someone who might know the person in Omaha. Once a participant received the folder they were to fill out the roster, detach one postcard and mail it back to the university (postage paid), if they knew the target person on a first name basis mail it directly to them, if not they were to mail the package to someone they thought might have a better chance of knowing the target, but they must know the person on a first name basis. Roughly half of the letters made it back to the university, the smallest distance was two, while some had nearly a dozen but the median was 5.5, rounded up to 6 (Milgram 1967). Obviously if the initial recipient and the target were on a first name basis with one another it was a strong tie between them. However, it was the weak ties between groups that ultimately enabled the letter to reach its target.

The strength of weak ties was further examined in *Job Search and Network Composition: Implications of the Strength-Of-Weak-Ties Hypothesis (Montgomery 1992)*. In this study Montgomery interviewed business executives to determine how they found their current job. They were asked if close friends and family or an acquaintance that helped them find their job. He concluded that managerial workers found their current job through weak ties 27.8 percent of the time. He determined this was because close friends and families were likely to have access to the same information and in order to gain access to outside information they had to utilize weak ties.

b. Translation

Much work has been done to translate routing techniques of social networks to computer and communications networks, particularly Delay Tolerant Networks. Delay Tolerant Networks are ideal for translating and implementing

social networking routing schemes due to their dynamic and volatile nature. With DTNs nodes are continuously joining and leaving the network. Additionally their respective link quality can vary throughout the life of the link. It is from these dynamic (weak) links that DTN's and mesh networks draw their strength. In traditional networks a link is relatively constant and always available for use unless it becomes too congested. Traditional networking algorithms like Open Path Shortest First (OPSF) route their traffic similar to the routing in the Milgram experiment, with the exception that routers advertise with whom they have connections to in order to route packets more efficiently. It isn't hard to think of packet switched network as a social network at all. Routers, and in some cases switches, can be equated to popular people in social networks. Along those lines the more friends a person has the more connections he or she is able to make. Likewise the more links a node has the more likely they are to form more links based on the fact that they are more likely to "know" the destination for a given message. It's slightly harder to correlate this example to DTN's because the topology of the network is constantly changing. Each node is likely to be in motion and therefore the nodes it is in contact with at any given time will change as well. At first glance the history of the contact may seem meaningless, but when collected, time predictive routing tables can be computed. It should be noted that these tables shouldn't be appended to indefinitely at the risk of being populated with contacts that are only made once and ultimately becoming unusable.

The authors of *Know Thy Neighbor: Towards Optimal Mapping of Contacts to Social Graphs for DTN Routing* analyze and evaluate several DTN routing protocols based on social networks to map the nodes of a network based on different criteria. The paper evaluates SimBet and BubbleRap under several contact generation models and measure the performance of each protocol (Granovetter 1973). Two approaches to time window based aggregations of contact, growing Time Window and sliding time window. Growing Time Window as contained in the original SimBet betweeness and similarity are calculated over

a social graph where there is an edge between two node if there was at least one contact at any time in the past. Betweeness of a node is defined as the fraction of the shortest paths between each possible pair of nodes going through this node. In sliding time window a limited time window is used for the centrality value calculations in BubbleRap (Pan, Crowcroft and Eiko 2011). Essentially, the time is split into 6 hour windows and only the contacts in the last six hours form the edges of the graph. The sliding time window approach helps to regulate the graph to regular contacts and eliminate random contacts.

Another DTN routing Protocol is Socially Selfish Aware Routing (SSAR) (Qinghua, Sencun and Guohong 2010). In SSAR the willingness of a node to route packets for another node is based on the social tie between them. Furthermore, nodes will forward the highest priority packets first in order to increase their "social status" (Ashby 1962). While it certainly has its merits, SSAR may not prove viable in a space based scenario based on the fact that users may be required to manually assign values for "stranger" nodes via an interface. This would be difficult to do for a satellite constellation because of the limited access to the satellite from the ground for programming and configuration changes.

c. Application

The traditional approach to satellite communications has been and currently remains to be the "bent pipe" approach. In other words, signals are unprocessed and only transponder. This is an inefficient approach to routing data. Consider the following scenario: A user on the ground sends a message to a distant user, because there is no terrestrial infrastructure connection the two users, satellites are used to relay the message. The message goes from the user on the ground to the satellite where it is then multiplexed and transmitted to the ground where the signal is then de-multiplexed, the ground station then determines where it needs to go and could possibly send it on to its final destination via terrestrial infrastructure, but this is not always the case. Often, the message is then multiplexed again at the ground station and sent back up to the

satellite to be retransmitted to its final destination. The reason behind this approach is cost and risk driven. Keeping the routing capability on the ground reduces the complexity of the design of the satellite, reduces the risk of having an on orbit failure rendering the satellite useless, an ultimately reduces the cost of the satellite. However, with advances in software definable radios, routing protocols and robustness of design, coupled with the ever-growing demand for access to large amounts of data by even the most disadvantaged user, the need to transition to packet switched satellite networks is becoming apparent.

The relatively long propagation times, coupled with other satellites constantly coming in and out of view requires a robust delay tolerant protocol for a packet switched satellite constellation to be viable. There are a number of approaches that can be utilized to provide worldwide coverage as previously discussed.

Examining the low earth orbit (LEO) approach through the lens of social mapping of DTNs exposes some familiarities already discussed. Satellites in LEO orbit the earth at tremendous speeds, about 7.5 km/s on average. Of course this speed will vary depending on the altitude of the satellite. Satellites that are in the same plane have fixed positions relative to one another and therefore the connections between these satellites can be classified and strong ties. The tremendous speed of the satellites becomes a significant problem as satellites move above 60 degrees north latitude and below 60 degrees south latitude. Here the relative velocities between satellites create a Doppler shift in the transmit and receive frequencies so great that the radios can’t compensate for it and the link disappears. Because these links between planes are not constant, these links are classified as weak ties. However, as demonstrated in both Granovetter and Montgomery’s works, the strength of the network resides in the weak ties. Without the weak ties between satellites in different planes, routing would not be possible without the use of a significant buffer for a store-carry-forward approach or the use of a ground station which is counterintuitive to a packet switched satellite network. Furthermore the links between users on the ground and the

satellites are weak ties because the satellites are in constant motion and the user switches from one satellite to another as they come in and out of view. If geosynchronous (GEO) satellites are included in this network the connections between the GEO satellites and the LEO satellites can be classified as weak ties as well based on the fact that the connections are not persistent. Again, the strength of the network lies in the weak ties.

Coupling the strong and weak ties of a routed satellite network with a sliding time window as stated in *Socially Selfish Delay Tolerant Networks* the network now has the ability to predict links before they become available to increase the routing efficiency of a DTN (Qinghua, Sencun and Guohong 2010).

To date there is only one satellite on orbit that is capable of performing routing onboard the satellite. INTELSAT 14 is a geosynchronous satellite that has a hosted payload, Internet Router In Space (IRIS) on board. The inclusion of an Internet router coupled with an IP modem on a commercial communications satellite allows for cross beam and cross band communications that are not seen in satellite communications, at least not without bouncing between the satellite and a ground station a few times. Because there is only one IP enabled satellite on orbit a true packet switched satellite constellation network does not yet exist.

B. HYBRID ROUTING PROTOCOL FOR SPACE

1. Overview

In order to more efficiently route network traffic through the unique space environment a new protocol is proposed, Hybrid Routing Protocol for Space (HRPS). HRPS is an adaptation based, hybrid routing protocol that leverages the unique predictability of orbital mechanics for use in a space-based network. The protocol is comprised of three layers a proactive, a reactive and a priority based decision matrix. These layers seek to maximize efficiency through the use of node location knowledge and minimize overhead required to maintain routes.

2. Tier One: Proactive

The Proactive layer is based on the premise that satellite orbits are well defined and the position of a satellite, and therefore the node itself, is well known at any given point in time. This can be predicted based on both the previous position and velocity data and verified through GPS or other telemetry measurements. The proactive layer takes a given satellite constellation and breaks it into a series of well-defined snapshots over time. These snapshots, or frames, are the basis for the predictability of Tier 1. For instance, a GEO satellite will see a LEO satellite come into view at a given reoccurring time. This LEO satellite will remain in view for just under half its orbit, and then break contact. This period of time that it is in view from first contact to break would be a frame with respect to that interface between GEO and LEO. Each type of interface would have a different frame period based on position and orbit of nodes.

The goal of this layer is to provide the nodes with an expectation at any given point in time of what other nodes should be within range and position to establish a link based solely on the orbital dynamics. From the basic orbital data, a Satellite Network Operating Center (SNOC) would periodically update the satellite constellation information to the satellites as required. This basic database is the Predetermined Satellite Database (PSD) that would reside on each node. From this, each node will send a Hello message to other nodes as they are expected to come into view as per the PSD. This Hello message will contain the Position (Pos), Period or orbit (Per), Time that further Hello message should be sent for that interface (Th/i) and a time stamp (Ts). Anytime a node receives a Hello message it will reply with an Acknowledgement (ACK) message containing Time received (Tr), Ts and Th/i for that node.

When a node receives a Hello message it stores the received information and uses it to update the PSD database, thereby correcting the satellite node positions due to any orbital effects and then computes a new routing table. When an ACK message is received by the original sender, it assigns a Trust Level (TL) value to the node it receives the message from. For each ACK message received

this value is increased by one until a value of five is reached in which case further Hello messages are not required and the proactive algorithm backs off. The TL levels range from 0–5, with zero meaning no familiarity and five, full trust. If at anytime a Hello message is sent, but no ACK message received, then the TL goes back to zero. As well, when a Hello message is sent over an interface that has a TL of zero or one, then the Hello message will also include the TL, Pos and Ts for all one-hop neighbors. It should be noted that this discovery process is based on the interfaces in the topology, such that each node has an interface for all neighbors.

The use of the Trust Levels serves two purposes. One, as knowledge and predictability of an interface increases, the need for the discovery overhead is reduced. If an ACK message is not received then, the discovery process is re-initiated and the neighbor information can be used to help discover the topology. Two, the Trust Levels will be used in Tier 3 for further decision making.

This Proactive portion of the HRPS protocol is a continual network discovery process that occurs, regardless of whether a message will be sent. It seeks to leverage the predictability of orbital mechanics in order to reduce the amount of overhead for network discovery. This process is capture in Figure 77.

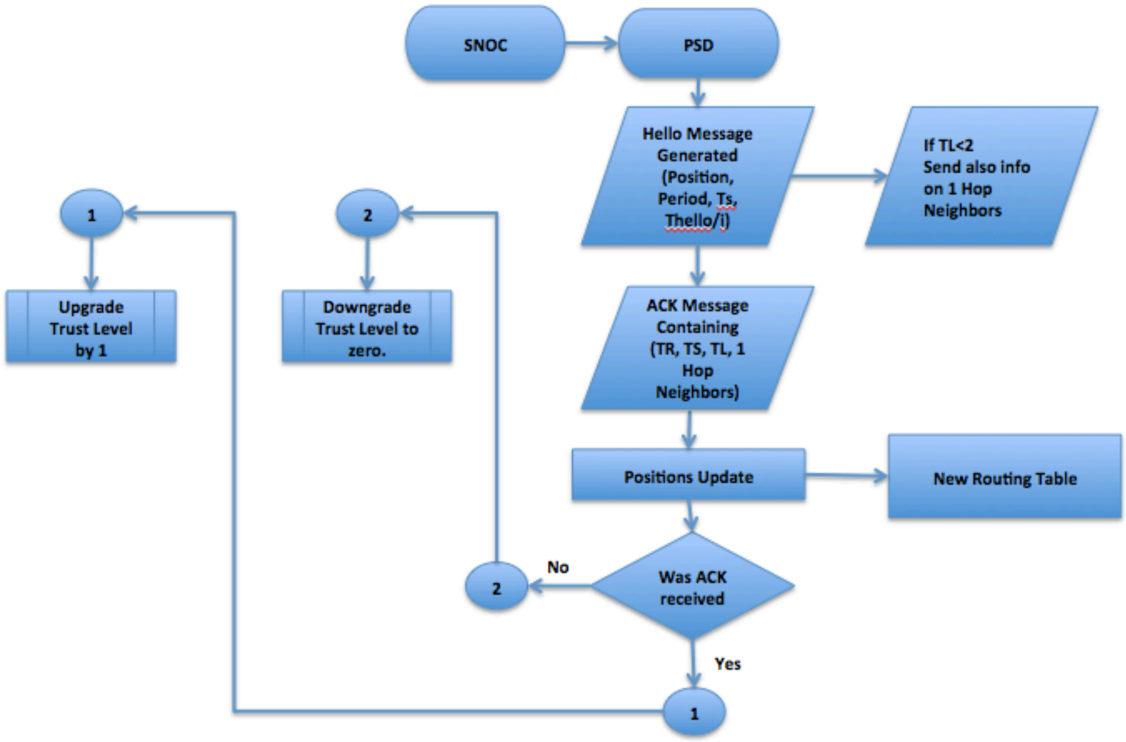

Figure 77. HRPS Tier One

3. Tier 2: Reactive

The Reactive layer is the second tier in the routing decision process. In most reactive MANET protocols, the nodes flood the network with *route request* packets in order to establish links with neighbors before traffic can be sent. Instead of sending route requests for all possible nodes in the network, the second tier sends requests to all expected neighbors based on information passed from the first layer. In other words, in each frame, the reactive layer will search for neighbors based on expected positions and establish links with those it can receive. Because of variability in the space environment, satellite condition and other factors, a predicted link may not actually exist. Route request responses will be noted with link metrics for signal strength, data rate and delay.

The Reactive layer begins its process when a message is to be sent. The first step is to use the routing table from Tier 1 to choose the three shortest routes that could be utilized to transmit the message to its destination. At this

point the sending node will send a Link State Request (LSREQ) to the destination across these paths. In response to this request, either the destination node or one of its neighbors who has the appropriate information will send a Link State Acknowledgement (LSACK) message with the following information for the interface paths: Bit Error Rate, Latency, Available Bandwidth and TL. This data is stored on the space routers in the form of a Management Information Base (MIB) and will be further discussed in a follow-on section. With the link-state information in hand, the algorithm will determine the best route for the packets to be forwarded.

Due to varying conditions within the space environment and changes in traffic loads across the network, the Reactive part of the protocol allows more advanced routing decisions. By utilizing the link-state metrics, HRPS becomes an adaptive protocol based on changing network dynamics in its route decision process. This could even extend to altering routes to take better advantage of weak ties or adjusting radio power or bandwidth as required. This process is capture in Figure 78.

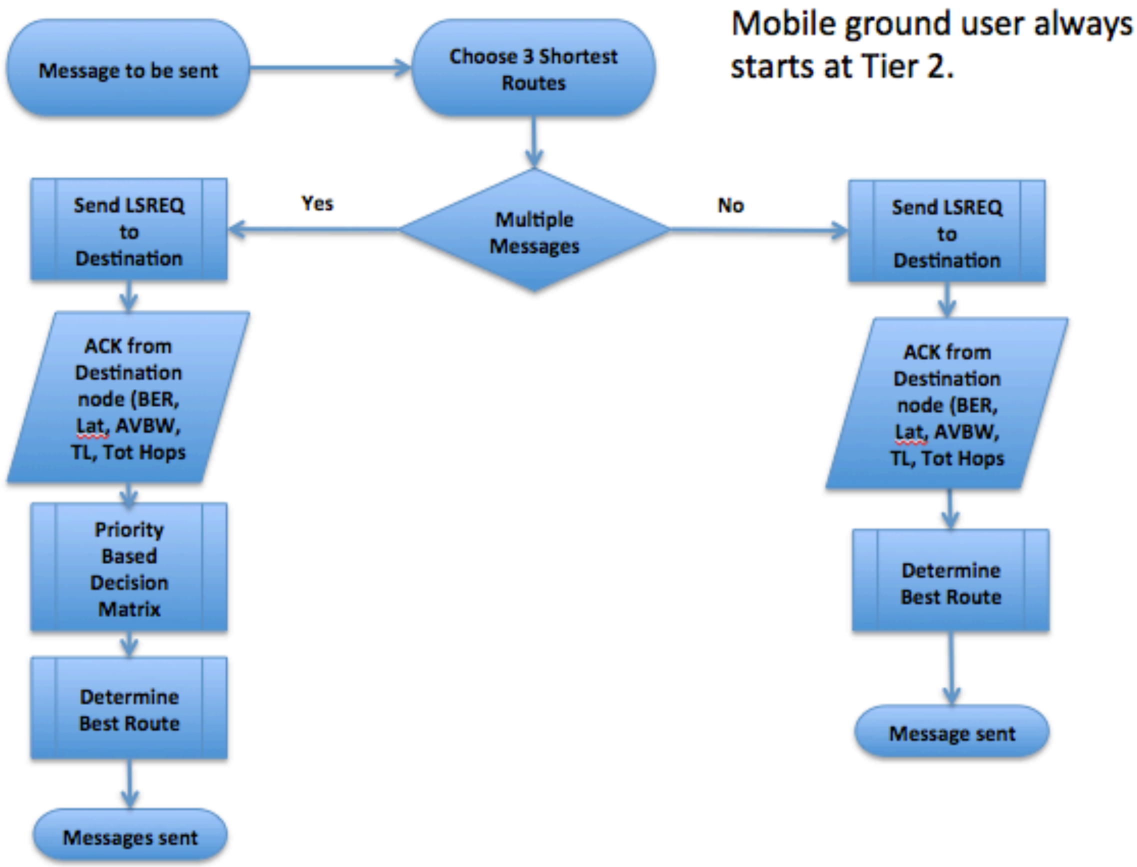

Figure 78. HRPS Tier 3

4. Tier 3

The final layer in HRPS is the Priority Based Decision Matrix. This is a decision matrix that has been predefined in order to pass traffic over the most appropriate path. Because this protocol is designed for military use it will be described in that vein, but can easily be altered for commercial use. When multiple messages are to be sent at the same time, the third tier will be used to route the traffic based on a predetermined value system. This is done in order to use the links most efficiently based on priority of the message. The decision matrix will take into account the priority of the message and then choose the link based on the link metrics discovered in tier two. For instance a high priority message for a Casualty Evacuation would be assigned the most robust link and route, as compared to a request for a routine weather update. The decisions

matrix can be setup to choose priority based on user, message type, or other criteria. This element will operate at the Application layer as compared to the other tiers which operate at the Network layer.

5. Space MIB

In order to fully realize the potential of HRPS, a network management model must be introduced in order to allow remote monitoring and configuration of nodes on the network. This is commonplace in most networks today. For the proposed network configuration, the Simple Network Management Protocol (SNMP) is well suited for this requirement. The reason for this is that SNMP uses a Management Information Base (MIB) to store and exchange management information. There are a wide range of MIB types and sizes based on the requirements of the network and management model as defined in Internet Engineering Task Force (IETF) Request For Comment (RFC) documents. (Subramanian 2000).

MIBs are stored as a virtual information tree on each node. The tree is broken into branching subgroups with managed objects within each subgroup. These pieces of data are collected and stored by the router and can be accessed by a NOC, other node, or an application. In the case of the space network, the MIB can be update to include the variables required for the HRPS protocol messages. Items such as the position, period and others can be pulled for the GPS unit, telemetry feed, and other sources and stored in the MIB. This allows network managers to access and update the MIB variables through SNMP as required to maintain the network as well. Other items not required by the HRPS protocol can also be included in the Space MIB in order to provide better diagnostics for the network health. The following table shows recommended variable for Space MIB.

MIB Variable
Node Position, Period, Th/I, Ts, Tr, TL
Satellite Battery State of Charge
Temperature of Transmitter
Single Event Latchups
Single Event Upsets

Table 11. Space MIB Variables

6. HRPS Summary

The HRPS protocol with Space MIB was designed in order to take advantage of the uniquely predictable nature of a space based mobile ad-hoc network. As noted, the size, distance and extreme mobility of this environment is not well suited to the way traditional MANET protocols are designed. By combining Proactive and Reactive elements, this protocol seeks to maximize performance and reliability in this environment. The process of using Weak Tie links as there are available increases the strength of the network and allows for a large number of routes available as the topology changes as opposed to only utilizing fixed pipes in more traditional space networks. In the Reactive Tier, aspects of adaptation borrowed from social networking theory allow the protocol to receive feedback and adjust routing

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Based on the four satellite constellations that were designed and modeled, it was found that all could provide the required coverage. The Hosted and Hybrid SCNMs contained the most links and were therefore the most robust due to the Weak Tie advantage. With respect to cost, the Hybrid model would be the most expensive and the Hosted the most economical. In the near term, the Hosted model would be the most likely candidate to begin forming such a satellite network.

In the case of the network modeling, no single protocol or configuration stood out above all others. It is possible that analyzing the models over a longer simulation time period may show a greater divergence between the protocols. It was found however that the LEO and Hybrid models had a shorter relative latency, which can be used as an advantage with time critical message traffic. Because no single protocol was especially well suited to the Space Networks tested, a new protocol concept was developed. Based on the predictability of orbital mechanics and utilizing social networking theories, the Hybrid Routing Protocol for Space is designed to more efficiently route traffic in this unique MANET network.

B. RECOMMENDATIONS

In order to further develop the HRPS protocol it will need to be coded for further testing in the lab environment. After this is accomplished it can be compared to other standing protocols and adjusted for efficiency. The authors also recommend further testing of HRPS and other MANET protocols over the IS-14 IRIS framework to further study the protocol development considerations.

With respect to the use of the various constellations in the development of future space networks, it is recommended that hosted payloads be considered.

Whether adding the routing capability to the next generation Iridium satellites, or including as secondary payloads on future satellite launches as proposed, the hosted method has been shown as an economical way to increase space networking infrastructure. In time, these capabilities can be added to small sats and cubesats using small form factor routing capabilities.

C. RESEARCH QUESTIONS

The goal of this thesis was to develop, model and analyze a space based communications network capable of providing the performance, coverage and logistics required of current and future military needs. In order to accomplish this, four satellite constellations were modeled to provide required coverage and compared for utility. These models found that while all are capable of providing the required services, the Hosted model is the most likely near term solution. By modeling these constellations with on-board routing capability, satellite constellation network models were developed and analyzed using current MANET routing protocols. Based on the results from thesis simulations a new protocol is proposed, designed specifically for the unique space environment.

D. AREAS FOR FURTHER RESEARCH

One area of future research is a cost benefit analysis of a hosted payload approach compared to a dedicated spacecraft approach (designing an entirely new satellite). This would provide an interesting comparison between the two approaches and determine which approach is a more efficient use of government resources.

LIST OF REFERENCES

Abedi, Saied, and Sunil Vadgama. “A Radio Aware Random Interative Scheduling Technique for High Speed Downlink Packet Access.” *Vehicular Technoloby Conference.* VTC 2002-Fall: IEEE, 2002. 2322–2326.

Adams, W. S., and L. Rider. 1987 “Circular Polar Constellations Providing Continuous Single or Multiple Coverage Above a Specified Latitude.” *The Journal of Astronautical Sciences* (American Astronautical Society) 35, no. 2: 155–192.

Analytical Graphics Inc. *STK Product webpage.* 2012. https://www.agi.com/products/by-product-type/applications/stk/Default.aspx (accessed May 28, 2012).

Ashby, W. Ross. 1962 “Principles of the Self-Organizing System.” *University of Illinois Symposium*,: 255–278.

Ashby, W. Ross, and Roger C. Conant. 1970: “Every Good Regulator of a System Must be a Model of that System.” *International Journal of Systems Science* 1, no. 2 : 89–97.

Burch, Ron. 2011 “The Case For Disaggregation Of U.S. MILSATCOM.” *Military Communications Comference-Track 6-Department of Defense Programs.* :2286–2291.

Chakravarthy, Balaji S. 1982 “Adaptation: A Promising Metaphor for Strategic Management.” *The Academy of Management Review* 7, no. 1: 35–44.

Chung, Junho, et al. 2008 “A Route Recovery Scheme in Dymo For Efficient Ad-oc.” *International Conference on Ubiquitous City Technology.* Sogang University, .

Clarke, Arthur C. 1945 “Extra-Terrestrial Relays: Can Rocket Stations Give World-Wide Radio Coverage?” *Wireless World* 305–308.

Comer, Douglas. 2004 *Compter Networks and Internets.* West Lafayette: Pearson Prentice Hall.

Corson, M. Scott, and Joseph Macker. 1999 “MANET Routing Protocol Performance Issues and Evaluation Considerations.” *The Internet Society* RFC 2501

Cuevas, Enrique G., H. Warren Kim, Hadi Esiely-Barrera, Zhuangbo Tang, and Brad Ward. 2010 *Internet Routing in Space Joint Capability Technology Demonstration.* Operational Utility Assessment, Johns Hopkins University.

Curtis, Howard. *Orbital Mechanics for Engineering Students.* Second. 2009 San Diego, CA: Elsevier.

Davis, Paul K., and James Bigelow. 2010 *Experiments in Multiresolution Modeling.* DARPA, Rand Corporation, Santa Monica: RAND Corporation, 1998.

Dean, Tamara. *Network+ Guide to Networks.* Boston, MA: Course Technology, Cengage Learning.

Du, Hongfei. 2008 "Proportional Channel-Aware Packet Scheduling in Satellite Multimedia Broadcasting." *Vehicular Technology Conference*.

Ekberg, Jan-Erik. 2008 "Ad-Hoc Routing Protocols: The STAR routing protocol."

Elbert, Bruce R. 2008 *Introduction to Satellite Communication.* Norwood: Artech House.

Freeman, Linton C. 1979 "Centrality in Social Networks Cenceptual Clarification." *Social Networks* 1, no. 3: 215–239.

Fujimoto, Richard M., Kalyan S. Perumalla, and George F. Riley. 2007 "Network Simulation." Morgan and Claypool 2007.

Granovetter, Mark S. 1973 "The Strength of Weak Ties." *American Journal of Sociology* 78, no. 6: 1360–1380.

Hess, Phill. 2012 *Glossary of Terms.* 2 14, 2012. asd-www.larc.nasa.gov/instrument/glossary.k.html (accessed August 23, 2012).

Holland, John H. 1962 "Outline for a Logical Theory of Adaptive Systems." *Journal of the ACM*: 297–314.

Hopfield, J. J. 1982 "Neural Networks and Physical Systems with Emergent Collective Computational Abilities." *Proceedings of the National Academy of Sciences* 79: 2554–2558.

Hossmann, T., T. Spyropoulos, and F. Legendre. 2010 "Know Thy Neighbor: Towards Optimal Mapping of Contacts to Social Graphs for DTN Routing." *INFOCOM Proceedings.* IEEE, 2010. 1–9.

Huhtonen, Aleksandr. 2004 *Comparing AODV and OLSR Routing Protocols.* Telecommunication Software and Multimedia Library, Helsinki Unitversity of Technology, Aalto: HUT T-110.551 Seminar on Internetworking, 2004.

Inmarsat. 2012 *Our satellites-Inmarsat.* January 1, 2012. www.inmarsat.com/About/Our_satellites/The_Inmarsat-4axpz?language=EN&textonly=false (accessed September 2, 2012).

Ippolito Jr., Louis J. 2008 *Satellite Communications Systems Engineering.* West Sussex: Wiley, 2008.

Iridium. *Iridium.* 1 1, 2012. www.iridium.com/About/Hosted Payloads.aspx (accessed September 1, 2012).

Ivancic, William, et al. 2005 *Secure, Network-Centric Operations of a Space-Based Asset: Cisco Router in Low Earth Orbit (CLEO) and Virtual Mission Operations Center (VMOC) .* Tech Memo, Scientific and Technical Information Program Office, National Aeronatics and Space Administration, Glenn Research Center: NASA, 2005.

Johns Hopkins University Applied Physics Laboratory. 2010 *Internet Protocol Routing In Space Joint Capabilites Technology Demonstration (IRIS JCTD) Operational Utility Assesment.* OUA, Applied Physics Labratory, Johns Hopkins University, Space and Missile Defense Battle Lab, 2010.

Kohonen, Teuvo. 1990 “The Self-Organizing Map.” *Proceedings of the IEEE* 78, no. 9 (Sep 1990).

Lang, Daniel. 2003 *A comprehensive overview about selected Ad Hoc Networking Routing Protocols .* Department of Computer Science, Munchen: Technische University at Munchen , 2003.

Michaelis, Anthony R. 1965 *From Semaphore to Satellite.* International Telecommunication Union, 1965.

Milgram, Stanley. 1967 “The Small World Problem.” *Psychology Today*, May 1, 1967: 61–67.

Montgomery, James D. 1992 “Job Search and Network Composition: Implications of the Strength of Weak Ties Hypothosis.” *American Sociological Review* 57, no. 5 (1992): 586–596.

Neuchterlein, J, and P Weiser. 2005 *Digital Crossroads: American Telecommunications Policy in the Digital Age.* Cambridge, Massachusetts: The MIT Press, 2005.

Newman, M. E. J. 2001 “Clustering and Preferenctial Attachment in Growing Networks.” *Physical Review* (American Physical Society) 64 (July 2001).

Orbital. 2012 “Hosted Payloads.” *Orbital.* January 1, 2012. www.orbital.com/hostedpayloads/ (accessed September 6, 2012).

Pan, Hui, Jon Crowcroft, and Yoneki Eiko. 2011 “BUBBLE Rap: Social-Baed Forwarding in Delay Tolerant Networks.” *IEEE Transactions on Mobile Computing*, 2011: 1576–1589.

Pratt, S.R, R.A. Raines, Fossa C.E., and M.A. Temple. 1999 “An Operation and Performance Overview of the Iridium Low Earth Orbit Satellite System.” *Communications Surveys and Tutorials*, April 1, 1999: 2–10.

Qinghua, Li, Zhu Sencun, and Cao Guohong. 2010 “Routing in Socialy Selfish Delay Tolerant Networks.” *INFOCOM.* IEEE, 2010. 1–9.

Qualnet. 2011 *QualNet 5.2 Developer Model Library .* Los Angeles: Scalable Network Technologies, Inc., 2011.

Qualnet. 2011 *QualNet 5.2 Model Library Index .* Los Angeles: Scalable Network Technologies, Inc., 2011.

Qualnet. 2011 *Qualnet 5.2 Wireless Model Library.* Los Angeles: Scalable Network Technologies, Inc., 2011, 204.

Qualnet. 2011 *Qualnet Programmer’s Guide.* Los Angeles: Scalable Network Technologies, Inc., 2011.

Rainey, Larry B. 2004 “Space Modeling and Simulation.” (Aerospace Press) 2004.

Ramaprasad, Arkalgud. 1983 “On the Definition of Feedback.” *Behavioral Science* 28 (1983): 4–13.

Skjernvold, Espen, Trude Hafsoe, Frank T. Johnsen, and Ketil Lund. 2009 “Delay and Disruption Tolerant Web Services for Heterogeneous Networks.” *IEEE*, 2009: Paper ID# 900190.

Subramanian, Mani. 2000 *Network Management: Principles and Practice.* Reading, MA: Addison Wesley Longman, Inc, 2000.

TerreStar. 2012 *www.terrestar.* January 1, 2012. www.terrestar (accessed Febuary 1, 2012).

Wasserman, Stanley.1994 “Social Network Analysis in the Social and Behabioral Sciences.” *Social Network Analysis: Methods and Applications*, 1994: 1–27.

Wertz, James, and Wiley Larson. 2007 *Space Mission Analysis and Design.* Hawthorne, CA: Microcosm, 2007.

Wertz, James, David Everet, and Jeffery Puschell. 2011 *Space Mission Engineering: The New SMAD.* Hawthorne, CA: Microcosm, 2011.

Wood, Lloyd. 2001 “Internetworking with satellite constellations.” PhD Thesis, School of Electronics, Computing and Mathematics, University of Surrey, 2001.

Wood, Lloyd. 1995 “Network performance of non-geostationary constellations equipped with intersatellite links .” Masters Thesis, Centre for Satellite Engineering Research, University of Surrey, 1995.

Wood, Lloyd, et al. 2006 “Adopting Internet Standards for Orbital Use.” *19th Annual AIAA/USU Conference on Small Satellites.* 2006.

Wood, Lloyd, George Pavlou, and Barry Evans. 2001 “Effects on TCP of Routing Strategies in Satellite Constellations.” *IEEE Communications Magazine*, March 2001.

Wood, Lloyd, Wesley M. Eddy, Will Ivancic, Jim McKim, and Chris Jackson. 2007 “Saratoga: a Delay-Tolerant Networking convergence layer with efficient link utilization.” *Third International Workshop on Satellite and Space Communications.* 2007.

Wood, Lloyd, William Ivancic, Dave Stewart, Dan Shell, and Phil Paulsen. 2005 *Secure, Network-Centric Operations of a Space Based Asset.* Technical Memo, Glenn Research Center: National Aeronautics and Space Administration, 2005.

Xu, Kaixin, Ken Tang, Rajive Bagrodia, Mario Gerla, and Michael Bereschinsky. 2011 *Adaptive Bandwidth Management and QoS Provisioning in Large Scale Ad-Hoc Networks.* Computer Science Department, University of California, Los Angeles, Boston: IEEE, 2003.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

4. Director, Training and Education,
   MCCDC, Code C46
   Quantico, Virginia

5. Director, Marine Corps Research Center,
   MCCDC, Code C40RC
   Quantico, Virginia

6. Marine Corps Tactical Systems Support Activity
   (Attn: Operations Officer)
   Camp Pendleton, California

7. Head, Information Operations and Space Integration Branch,
   PLI/PP&O/HQMC,
   Washington, DC

8. Eric Watko,
   CISCO Space Division, Research Triangle Park,
   Raleigh-Durham, North Caroline